THE PROCEEDINGS OF THE PHYSICAL SOCIETY

VOL. 52, PART 3 1 May 1940 No. 291

A NON-CARTESIAN MIRROR, REFRACTING SURFACE AND LENS TREATMENT

BY ERIC J. IRONS, PH.D., F.INST.P.,
Queen Mary College, London (c/o King's College, Cambridge)

ABSTRACT. The paper contains a detailed account of a treatment of reflection and refraction based on a sign convention which involves the convergence and divergence of beams of light, and gives positive signs to the powers of concave mirrors and convex lenses situated in air. The essence of the treatment is given in four rules which are applied both to single and multiple reflecting and refracting surfaces, and substantiated by proofs which involve finite objects and both convergent and divergent incident light. The treatment includes a discussion of Newton's formula and the cardinal points of a lens system, and there are appended the solutions of ten problems to illustrate the use of the method proposed.

§1. INTRODUCTION

DURING the past decade the subject symbolized by "$1/f=1/v\pm1/u$" has received considerable attention from those who, by their standing in pedagogic and technical circles, are well qualified to discuss it. Nevertheless, unanimity as to the best convention is still lacking and, in fact, textbooks are now appearing in which problems are treated by different methods in parallel columns. These facts have emboldened the writer, who lays no claim to specialized knowledge, to record a treatment involving a non-Cartesian convention in the hope that, although not entirely original, it may be of interest to teachers of elementary students as a summary illustrated by examples.

§2. BASIS OF TREATMENT AND RULES FOR CALCULATIONS

As the idea of a real image is more easily comprehended than that of a virtual image, it is convenient to deal primarily with light which after reflection or refraction is rendered convergent, and to introduce the subject of image-formation by consideration of converging mirrors and lenses. Association of both convergent light and converging systems with a positive sign does not then appear altogether unnatural when examination of the light emitted by a body or the properties of diverging systems shows the necessity for some such differentiation. The

treatment described is therefore based upon the ideas of convergence and divergence and is summarized in the following rules :—

(1) The convergence of a beam with respect to a reflecting or refracting surface is measured in dioptres by $(\pm)$ the product of the refractive index of the medium in which it is travelling and the reciprocal of the distance expressed in metres between the surface and the point $\binom{\text{to}}{\text{from}}$ which it $\binom{\text{converges}}{\text{diverges}}$.

It will be noted that (i) the quantity by which the refractive index is multiplied is the curvature of the wave-front at the reflecting or refracting surface; (ii) a parallel beam has zero convergence; and (iii) divergence is to be regarded as negative convergence, and the algebraic sign is associated with a physical concept.

(2) The effect of a reflecting or refracting surface on a beam is expressed by the formula

Final convergence = initial convergence + (converging) power,

and if the final convergence is $(\pm)$ a $\binom{\text{real}}{\text{virtual}}$ image is formed.

(3) (*a*) The converging power of a surface is $\pm n_1/r \pm n_2/r$ dioptres, where n_1 and n_2 are the refractive indices of the media in which the incident and reflected or refracted beams travel, and r is the radius of the surface in metres: the $(\pm)$ sign is used when the surface is $\binom{\text{concave}}{\text{convex}}$ to the medium whose refractive index follows in the expression.

(*b*) When the distances between surfaces are inappreciable, the power of the system is the sum of the powers of the reflecting and refracting surfaces. We infer that (i) As light is reflected *from* a surface, $n_1 = n_2 = n$ (say), and $\binom{\text{concave}}{\text{convex}}$ mirrors have converging powers equal to $\pm 2n/r$, which simple geometry shows to be equivalent to $\pm 1/f$ in air. A plane mirror has zero converging power and merely changes the direction of the light. (ii) As light is refracted *through* a surface, $\binom{\text{concave}}{\text{convex}}$ surfaces have converging powers equal to $\mp(n_2 - n_1)/r$. (iii) The converging power of a thin double $\binom{\text{concave}}{\text{convex}}$ lens made of material of refractive index n_2 surrounded by a medium of refractive index n_1 is

$$\mp\frac{(n_2-n_1)}{r_1} \pm \frac{(n_1-n_2)}{r_2} = \mp(n_2-n_1)\left(\frac{1}{r_1}+\frac{1}{r_2}\right),$$

where r_1 and r_2 are the radii of the first and second surfaces.

(4) (*a*) The transverse magnification = (length of image) ÷ (length of small object perpendicular to the axis) = (initial convergence) ÷ (final convergence), and if these convergences are of the $\binom{\text{same}}{\text{opposite}}$ sign the image is $\binom{\text{erect}}{\text{inverted}}$.

(*b*) The longitudinal magnification=(length of image)÷(length of small object along the axis)=(ratio of refractive indices of media in which light travels after and before reflection or refraction respectively)×(initial convergence)^2÷(final convergence)^2. If the initial and final convergences are of the $\begin{pmatrix}\text{same}\\\text{opposite}\end{pmatrix}$ sign, the end of the object nearer the reflecting or refracting surface has its image $\begin{pmatrix}\text{nearer to}\\\text{more remote from}\end{pmatrix}$ that surface than the image of the other end.

The rules are here stated in a more general form than is necessary for beginners. Thus, for example, mention of refractive index and dioptres in rule 1 may be omitted at the discretion of the teacher, and rule 3 may be given in terms of the focal length f, thus: "The converging power of concave mirrors and convex lenses is $+1/f$; of convex mirrors and concave lenses $-1/f$": and use may be made of the alternative proofs for lenses given in § 3 (C') below, n_1 being assumed to be unity.

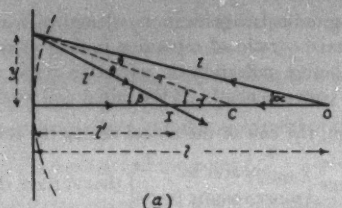
It should be noted that the signs of the powers are the same as those employed in the optical industry, and that the significance of the rules may be explained equally well by ray or wave-curvature treatments.

Moreover, it is hoped that by expressing rule 2 in words and showing that it summarizes the facts concerning the relative positions of objects and images, the confusion that sometimes arises between "the algebraic sign prefixed to a symbol in an equation and the positive or negative character of the numerical value which is assigned to a symbol when an actual numerical computation has to be made" will disappear.

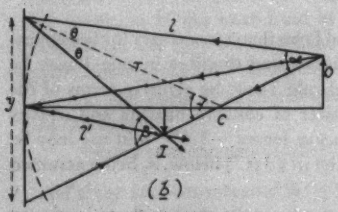
§ 3. PROOFS OF FORMULAE

In proving the formulae and substantiating the rules given in the last section it is, at least in the writer's opinion, definitely desirable to convince a pupil that the formulae are directly applicable to convergent as well as to divergent incident light, and thereby obviate the necessity for arguments based on the reversal of optical paths. Proofs of this fact are readily evolved, and examples are in consequence included. On the other hand, there appears to be little need to follow the practice of some texts and deal first with a point object on the axis and subsequently with an extended object perpendicular to the axis, especially as a common type of point-object proof may be modified so as to include extended objects and at the same time to emphasize that the formulae are true only for rays associated with small objects and incident nearly normally on surfaces of small curvature. In the typical proofs which follow an attempt has been made to keep the diagrams as similar as possible and to preserve uniformity of nomenclature.

(*A*) *Spherical mirrors*, figure 1. The reflecting and refracting surfaces with which we are concerned are supposed to be of such small curvature that continuous straight lines are used to represent their position, while their natures are indicated

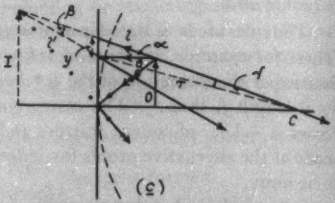


(a)



(b)

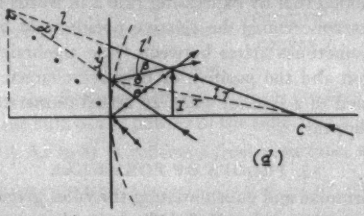
Object beyond focal plane. $2\theta=\beta-\alpha=2(\gamma-\alpha)$; $\beta=-\alpha+2\gamma$; $y/l'=-y/l+2y/r$; $1/l'=-1/l+1/f$.



(c)

Object between focal plane and pole. $2\theta=\alpha+\beta=2(\alpha-\gamma)$; $-\beta=-\alpha+2\gamma$; $-y/l'=-y/l+2y/r$; $-1/l'=-1/l+1/f$.

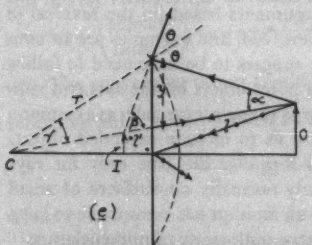
Incident light diverging.



(d)

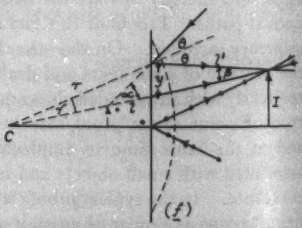
Incident light converging. $2\theta=\alpha+\beta=2(\beta-\gamma)$; $\beta=\alpha+2\gamma$; $y/l'=y/l+2y/r$; $1/l'=1/l+1/f$.

CONCAVE MIRROR.



(e)

Incident light diverging. $2\theta=\alpha+\beta=2(\alpha+\gamma)$; $-\beta=-\alpha-2\gamma$; $-y/l'=-y/l-2y/r$; $-1/l'=-1/l-1/f$.



(f)

Incident light converging. $2\theta=\alpha+\beta=2(\alpha-\gamma)$; $\beta=\alpha-2\gamma$; $y/l'=y/l-2y/r$; $1/l'=1/l-1/f$.

CONVEX MIRROR.

Figure 1.

by broken curved lines. Throughout the treatment vertical distances are for clarity exaggerated in comparison with horizontal distances.

In figure 1 (*a*) are shown the paths of two rays of light from an object on the axis before and after reflection at a concave mirror. One ray makes an angle θ with the mirror, the other passes through the centre of curvature, and it is easy to show that $\beta=-\alpha+2\gamma$. If the angles are sufficiently small they may be replaced by their tangents or their sines, and it follows that $1/l'=-1/l+2/r$.

If now we consider the extremity of a small object perpendicular to the axis of the mirror (figure 1 (*b*)), and draw from it two rays similar to those of figure 1(*a*), it will again be observed that

as $$2\theta=\beta-\alpha=2(\gamma-\alpha),$$

therefore $$\beta=-\alpha+2\gamma,$$

and $$1/l'=-1/l+2/r=-1/l+1/f\ *,$$

if the angles are so small that the distances of object and image from the mirror are sensibly the same whether measured along a ray or parallel to the axis. As this result is independent both of y and of the lengths of the small object and image it follows that, subject to the conditions already cited, all rays from a point on O pass through the corresponding point on I, and the image is a straight line perpendicular to the axis.

The remaining diagrams in figure 1 are self-explanatory, and in each instance it may be observed that the term on the left-hand side of the equations is the final convergence, the first term on the right-hand side the initial convergence, and the last term the converging power. To obtain these quantities in their more general form it is of course merely necessary to multiply throughout by the refractive index of the medium.

To verify rule 4 for transverse magnification it is convenient to show the path of that ray from a point on the object which is incident at the pole (full dotted line): it may then be observed that this magnification $=|l'/l|=|1/l\div 1/l'|=$(initial convergence)$\div$(final convergence), and only if these quantities are of the same sign is the image erect. The subject of longitudinal magnification is dealt with at (*D*) below.

Two special positions of the object, viz. at the centre of curvature and in the focal plane of a concave mirror, call for comment. In the first case incident rays from any point on a small object meet the mirror at practically the same small angle and are therefore deviated by the same amount to form an inverted image such that $l=l'=r$ or $1/l'=-1/l+2/r$; while in the second case the standard diagram with the usual notation indicates that $\theta=\gamma$ (these being the angles of an isosceles triangle) and $\alpha=\theta+\gamma=2\theta$, i.e. the reflected rays are parallel.

(*B*) *Spherical refracting surfaces*, figure 2. As before the position of the surface is indicated by a straight line and its nature by a curved broken line. By each diagram is an equation demonstrating the truth of rules 2 and 3 (*a*), and in each instance the magnification may be found by considering the path of the ray—

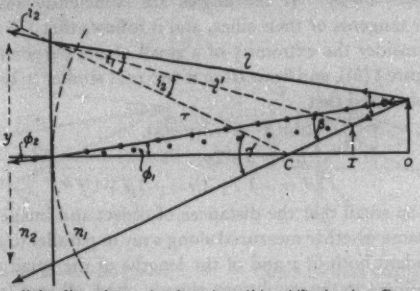
* In the note on power, at (*E*) below, r is shown to be equal to $2f$.

shown by the full dotted line—which is incident on the surface where it cuts the axis. This quantity may be seen to be given by

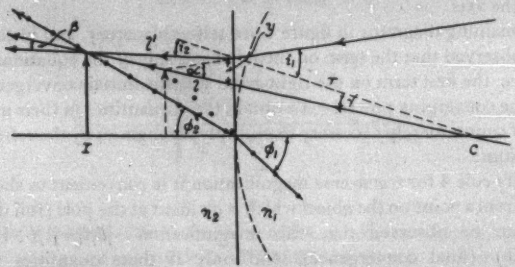
$$|l'\phi_2/l\phi_1| = |n_1/l \div n_2/l'|$$

$$= (\text{initial convergence}) \div (\text{final convergence}) ;$$

moreover, when the ratio is positive the image may be observed to be erect.

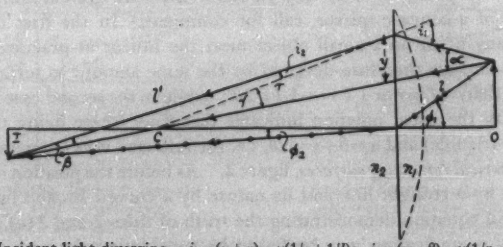


(*a*) Incident light diverging. $i_1=(\gamma-\alpha)=y(1/r-1/l)$; $i_2=(\gamma-\beta)=y(1/r-1/l')$; $n_1(1/r-1/l)=n_2(1/r-1/l')$; $-n_2/l'=-n_1/l-(n_2-n_1)/r$.



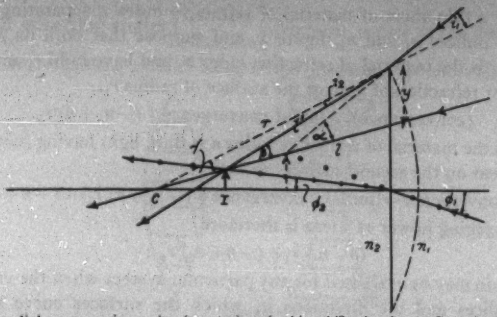
(*b*) Incident light converging. $i_1=(\gamma+\alpha)=y(1/r+1/l)$; $i_2=(\gamma+\beta)=y(1/r+1/l')$; $n_1(1/r+1/l)=n_2(1/r+1/l')$; $n_2/l'=n_1/l-(n_2-n_1)/r$.

CONCAVE SPHERICAL SURFACE.



(*c*) Incident light diverging. $i_1=(\gamma+\alpha)=y(1/r+1/l)$; $i_2=(\gamma-\beta)=y(1/r-1/l')$; $n_1(1/r+1/l)=n_2(1/r-1/l')$; $n_2/l'=-n_1/l+(n_2-n_1)/r$.

Figure 2.



(*d*) Incident light converging. $i_1=(-\gamma+\alpha)=\gamma(-1/r+1/l)$; $i_2=(-\gamma+\beta)=\gamma(-1/r+1/l')$; $n_1(-1/r+1/l)=n_2(-1/r+1/l')$; $n_2/l'=n_1/l+(n_2-n_1)/r$.

CONVEX SPHERICAL SURFACE.

For all examples, $n_1 \sin i_1=n_2 \sin i_2$; i. e., $n_1 i_1=n_2 i_2$ as angles are small; and one ray passes undeviated from the object through the centre of curvature of the surface.

Figure 2 (*cont.*).

The diagrams, which illustrate typical examples of refraction, are drawn on the assumption that $n_2>n_1$, but by supposing the direction of the light to be reversed, we may use the first and last two diagrams of figure 2 to show rays travelling in a denser medium and, incident respectively on convex and concave surfaces, refracted into a rarer medium. The relation between the quantities concerned still holds and may be rewritten to conform with our rules : thus, for the example illustrated in figure 2 (*d*),

$$-n_1/l=-n_2/l'-(n_1-n_2)/r.$$

When the small object for refraction is at the centre of curvature of a concave surface, rays from any point on it do not undergo appreciable refraction, so that a virtual image coincident with the object is formed, i.e.

$$l=l'=r \quad \text{and} \quad -n_2/l'=-n_1/l-(n_2-n_1)/r \text{ again.}$$

Rays of a convergent beam directed towards the centre of curvature of a convex surface are similarly unrefracted. It may be noted that light from an object between a concave surface and its centre of curvature would actually be *converged* and also that a convex surface *diverges* incident convergent light which, in the absence of refraction, would form a real image nearer its surface than its centre of curvature (figure 2 (*d*)). This does not imply that in such instances the power of the surface should be reversed for, as with mirrors, it is the effect on rays parallel to the axis that distinguishes converging from diverging systems.

(*C*) *Lenses*. The usual lens formulae are obtained readily and in their most general form by application of the results deduced in the preceding section. In courses for elementary students, however, consideration of refraction at a single spherical surface is normally omitted, and accordingly a more direct if restricted treatment of lenses is given at (*C'*) below.

Consider a thin sheet of material of refractive index n separating two media of refractive indices n_1 and n_2, figure 3, and suppose that both its surfaces are convex towards the material of refractive index n_1 and have radii r_1 and r_2 respectively. After refraction of light at the surface of radius r_1,

$$(\text{convergence})_1 = \text{initial convergence} + (-n_1+n)/r_1.$$

Inasmuch as the material of refractive index n is thin, light having $(\text{convergence})_1$ is now incident on the second surface, so that

$$\text{final convergence} = \{\text{initial convergence} + (-n_1+n)/r_1\} + (-n+n_2)/r_2$$

and the converging power of a lens is therefore

$$(n-n_1)/r_1 + (-n+n_2)/r_2 :$$

this expression may be evaluated for any particular system when the values of the refractive indices and the directions in which the surfaces curve are known.

Figure 3.

Moreover, as the

(length of the image that would have been formed by the first refraction had only that refraction occurred) ÷ (length of the object)

$$= (\text{initial convergence}) \div (\text{convergence})_1$$

and the

(length of the final image) ÷ (length of the first image)

$$= (\text{convergence})_1 \div (\text{final convergence}),$$

it follows that the

(length of the final image) ÷ (length of the object)

$$= (\text{initial convergence}) \div (\text{final convergence}),$$

and whether the image is erect or inverted depends on the sign of this quotient in the usual way.

Again, supposing light to diverge from an object distant l from the surface of radius r_1 and to form an image at a distance l' from the other surface, then

$$\pm n_2/l' = -n_1/l + [(n-n_1)/r_1 + (n_2-n)/r_2]$$

according as the image is real or virtual. Now when $l=\infty$, $l'=f_2$, where f_2 is the focal length of the lens for light transmitted to the medium of refractive index n_2: the power of the lens is therefore $\pm n_2/f_2$ and it may also, by a similar argument, be shown to be $\pm n_1/f_1$.

(*C'*) *Lenses; a direct treatment.* An alternative treatment for lenses surrounded on both sides by the same medium is based on two ideas :—

(1) A $\begin{pmatrix}\text{converging}\\ \text{diverging}\end{pmatrix}$ lens deviates light in the same way as a series of prisms

whose small refracting angles $\begin{pmatrix}\text{increase}\\ \text{decrease}\end{pmatrix}$ from the centre of the lens towards its periphery, the remaining parallel-sided strips producing no effect (figure 4)

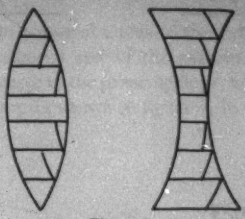


Figure 4.

(2) The deviation towards the base of a prism of a ray incident nearly normally from a medium of refractive index n_1 on a prism of small angle A whose material is of refractive index n is given by

$$\delta=(n/n_1-1)A,$$

and is therefore independent of the exact angle of incidence.

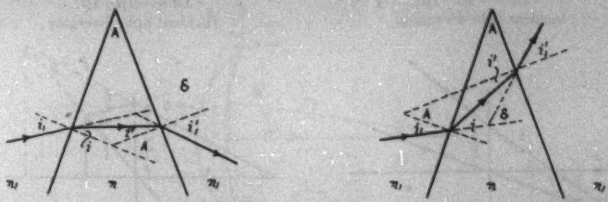
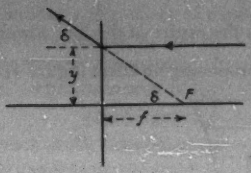


Figure 5.

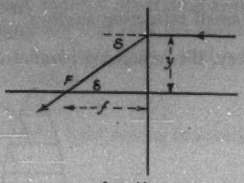
Proofs of this formula may be based on wave-front principles or on the ray treatment indicated in figure 5. Both results are summed up in the equation

$$\delta=\left(\frac{n-n_1}{n_1}\right)A$$

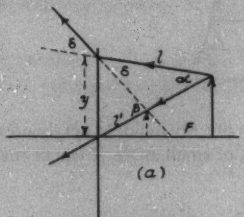
if we agree to interpret a negative value of δ as a deviation *away* from the base of the prism. From these two principles it follows that δ, the deviation of a ray produced at any point on a lens distant y from its centre, is independent of the angle of incidence, and figure 6 shows that, if a ray parallel to the axis of the lens is chosen, $\delta=y/f$. The diagrams of figure 6 then indicate the truth of the formulae of rules 2 and 4 (*a*) for divergent and convergent beams of light incident on converging and diverging lenses. It may be noted that in each instance we are supposing that the material of the lens has a greater refractive index than its



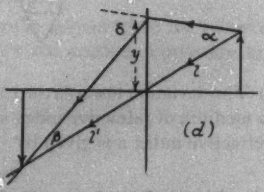
$\delta = y/f.$



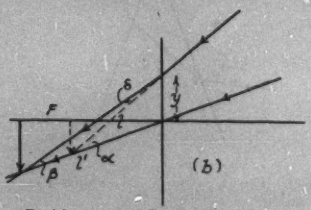
$\delta = y/f.$



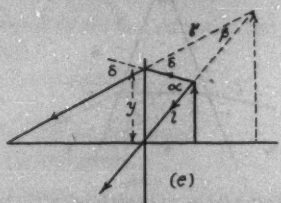
$\beta = \alpha + \delta$; $y/l' = y/l + y/f$;
$-1/l' = -1/l - 1/f$.
Incident light diverging.



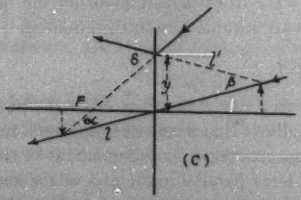
Real image. $\delta = \alpha + \beta$; $y/f = y/l + y/l'$;
$1/l' = -1/l + 1/f$.
Incident light diverging.



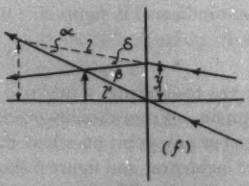
Real image. $\alpha = \beta + \delta$; $y/l = y/l' + y/f$;
$1/l' = 1/l - 1/f$.
Incident light converging.



Virtual image. $\alpha = \beta + \delta$; $y/l = y/l' + y/f$;
$-1/l' = -1/l + 1/f$.
Incident light diverging.



Virtual image. $\delta = \alpha + \beta$; $y/f = y/l + y/l'$;
$-1/l' = 1/l - 1/f$.
Incident light converging.
CONCAVE LENS.



$\beta = \alpha + \delta$; $y/l' = y/l + y/f$;
$1/l' = 1/l + 1/f$.
Incident light converging.
CONVEX LENS.

Figure 6.

surroundings, i.e. $n>n_1$, and that one ray of the incident beam is made to pass undeviated through the centre of the lens. To obtain the equations in their more general form it is, as in the case of mirrors, only necessary to multiply throughout by n_1.

It remains to relate the power of a lens to the radii of its surfaces and to the refractive indices of its material and of the surrounding medium. This may be effected by further recourse to the prism analogy, and the method of procedure is illustrated by some examples shown in figure 7, in which the lens thickness is

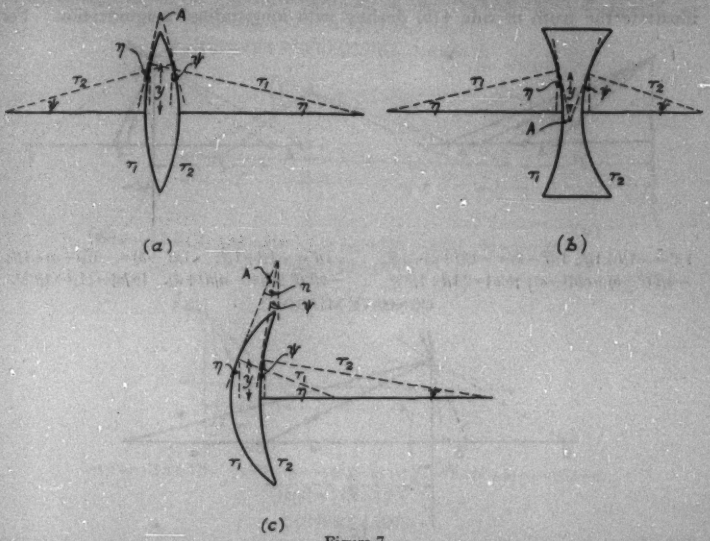


Figure 7.

greatly exaggerated. The effective angle A of the equivalent prism at a height y above the centre of the lens is $(\eta+\psi)$ in figures 7 (a) and 7 (b), and $(\eta-\psi)$ in figure 7 (c), but in each instance

$$\eta=y/r_1,\ \psi=y/r_2$$

and, corresponding to the conditions of figure 7 (c),

$$\delta=y/f=\left(\frac{n-n_1}{n_1}\right)A=\left(\frac{n-n_1}{n_1}\right)(y/r_1-y/r_2),$$

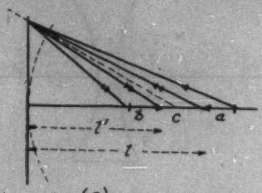
i.e. the converging power of the lens

$$(n-n_1)(1/r_1-1/r_2)\ \text{ is }\ n_1/f.$$

For any other combination of surfaces it is merely necessary to observe whether A is the sum or difference of the angles η and ψ and to preface the numerical value of the term $(1/r_1\pm1/r_2)$ with a $(\pm)$ sign according as the lens is $\begin{pmatrix}\text{thicker}\\ \text{thinner}\end{pmatrix}$ at the

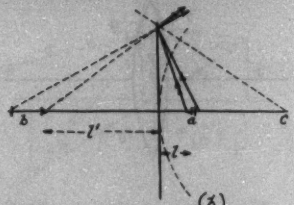
centre than at the periphery. This general equation has been developed so as to yield values of the power of a lens surrounded by air ($n_1=1$) and also by a material of greater refractive index, i.e. one for which $(n-n_1)$ is negative. In the latter circumstance the sign of the power changes and the result indicates that a convex lens of, say, water surrounded by glass is diverging. It is, however, contemplated that it will in general be sufficient for those elementary students for whom the method is appropriate to develop the argument with n_1 equal to 1 throughout.

(*D*) *Longitudinal magnification.* The diagrams of figure 8 are designed to illustrate the truth of rule 4(*b*) dealing with longitudinal magnification. For



(*a*)

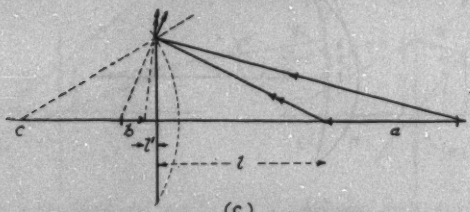
$1/l'=-1/l+1/f$; $1/(l'-b)=-1/(l+a)+1/f$;
$-b/l'(l'-b)=a/l(l+a)$; $|b/a|=\{1/l\div 1/l'\}^2$.



(*b*)

$-1/l'=-1/l+1/f$; $-1/(l'+b)=-1/(l+a)+1/f$;
$-b/l'(l'+b)=-a/l(l+a)$. $|b/a|=\{1/l\div 1/l'\}^2$.

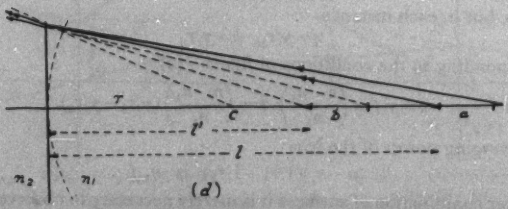
CONCAVE MIRROR.



(*c*)

$-1/l'=-1/l-1/f$; $-1/(l'+b)=-1/(l+a)-1/f$;
$-b/l'(l'+b)=-a/l(l+a)$; $|b/a|=\{1/l\div 1/l'\}^2$.

CONVEX MIRROR.

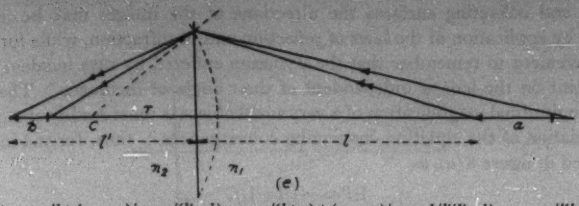


(*d*)

$-n_2/l'=-n_1/l-(n_2-n_1)/r$; $-n_2/(l'+b)=-n_1/(l+a)-(n_2-n_1)/r$; $n_2\,b/l'(l'+b)=-n_1\,a/l(l+a)$;
$|b/a|=(n_1/n_2)(l'/l)^2=(n_2/n_1)\{(n_1/l)/(n_2/l')\}^2$.

CONCAVE REFRACTING SURFACE.

Figure 8.

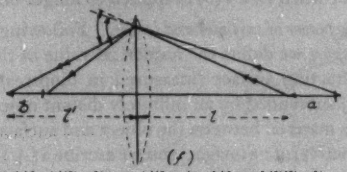


(e)

$n_2/l' = -n_1/l + (n_2 - n_1)/r$; $n_2/(l'-b) = -n_1/(l+a) + (n_2-n_1)/r$; $n_2 b/l'(l'-b) = -n_1 a/l(l+a)$;

$|b/a| = (n_1/n_2)(l'/l)^2 = (n_2/n_1)\{(n_1/l)/(n_2/l')\}^2$.

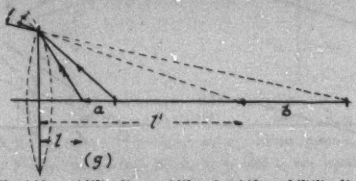
CONVEX REFRACTING SURFACE.



(f)

$1/l' = -1/l + 1/f$; $1/(l'-b) = -1/(l+a) + 1/f$; $-b/l'(l'-b) = -a/l(l+a)$;

$|b/a| = \{1/l \div 1/l'\}^2$.

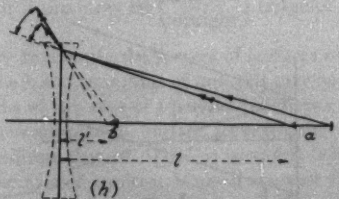


(g)

$-1/l' = -1/l + 1/f$; $-1/(l'+b) = -1/(l+a) + 1/f$; $-b/l'(l'+b) = -a/l(l+a)$;

$|b/a| = \{1/l \div 1/l'\}^2$.

CONVEX LENS.



(h)

$-1/l' = -1/l - 1/f$; $-1/(l'+b) = -1/(l+a) - 1/f$; $-b/l'(l'+b) = -a/l(l+a)$;

$|b/a| = \{1/l \div 1/l'\}^2$.

CONCAVE LENS.

In interpreting this figure it should be remembered that
(i) a and b are supposed to be small compared with l and l'.
(ii) A real image is associated with a converging pencil.

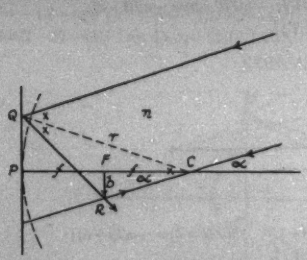
Figure 8 (*cont.*).

mirrors and refracting surfaces the directions of the images may be inferred directly by application of the laws of reflection and of refraction, while for lenses it is convenient to remember that the deviation suffered by rays incident at the same point on the lens is independent of their angle of incidence. The value of the longitudinal magnification of a very small object is most readily obtained by differentiation of the equation connecting convergences ; thus, for the instance illustrated in figure 8 (*a*), as

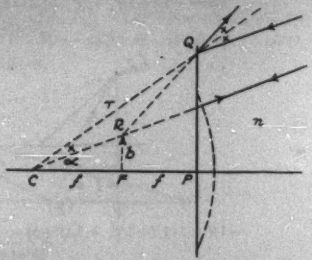
$$1/l' = -1/l + 1/f,$$
$$-\delta l'/(l')^2 = \delta l/(l)^2.$$

The more elementary if less elegant method shown in figure 8 indicates the magnitude of the error when rule 4 (*b*) is applied to longer objects.

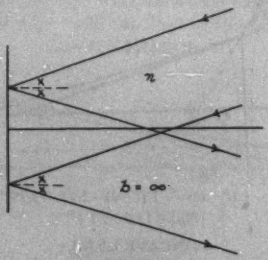
(*E*) *The converging power of surfaces and lenses.* Following the lead given in the Physical Society's *Report* we define the numerical value of the converging power of a reflecting or refracting surface (measured in dioptres) as the small angle (measured in radians) subtended by an infinitely distant object multiplied by the refractive index of the material between the object and surface and divided by the length (measured in metres) of its image ; and we ascribe a (±) power to a reflecting



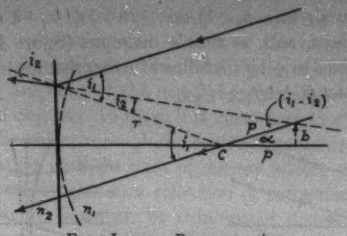
Inverted image. Power $= +n\alpha/b = +n/f = +2n/r$.
CONCAVE.



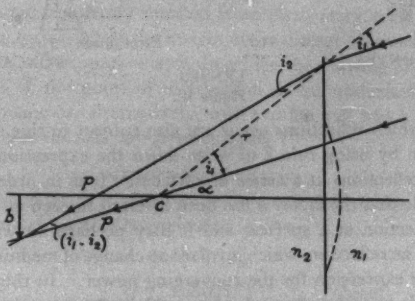
Erect image. Power $= -n\alpha/b = -n/f = -2n/r$.
CONVEX.



Power = 0.
PLANE.
MIRRORS.
Figure 9.



Erect Image. Power negative.
CONCAVE.



Inverted image. Power positive.
CONVEX.

$|\text{Power}| = n_1\alpha/b = n_1/p = n_1\{n_2/n_1 - 1\}/r = |(n_2 - n_1)/r|$, since $p/\sin i_2 = r/\sin(i_1 - i_2)$ and $r/p = (i_1 - i_2)/i_2 = (n_2 - n_1)/n_1$ as i_1 and i_2 are small.

REFRACTING SURFACES.

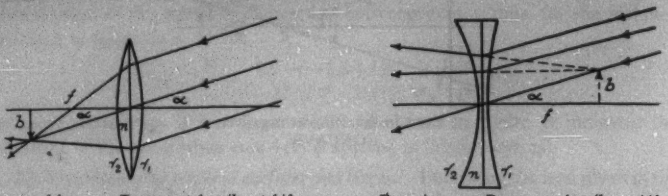
Figure 9 (*cont.*).

or refracting surface if it produces an $\begin{pmatrix}\text{inverted}\\ \text{upright}\end{pmatrix}$ image of an infinitely distant object.

It is now necessary to identify the powers of surfaces as defined above with those derived from rule 3, and an outline of how this may be effected is indicated in figure 9. The paths of two rays of a parallel pencil, one of which is directed towards the centre of curvature of the surface, serve to define the size of the image and for the mirrors indicate that, as ΔQRC is isosceles, $QR = RC$: for rays but slightly inclined to the axis of a mirror of small aperture this relation becomes $PF = FC$ and indicates that $2f = r$.

Now at the commencement of sub-section (*C*) above it is shown that the effective power of two nearly coincident refracting surfaces is given, with due regard to sign, by the sum of their separate powers; and it is clear that the same argument may be extended to any number of surfaces. As a particular instance of this general procedure we may note that the powers of the lenses shown in figure 10 when the

lenses are situated in air ($n_1=n_2=1$) are $\pm(n-1)(1/r_2+1/r_1)$, the sums of the powers of their surfaces, and, as simple diagrams (figure 10) indicate, are, by definition of power, equal to $\pm 1/f$: it follows further that in air the power of a series of thin lenses in contact is $\pm 1/f_1 \pm 1/f_2 \pm \ldots$



Inverted image. Power $=+1 \cdot \alpha/b=+1/f$. CONVEX.

Erect image. Power $=-1 \cdot \alpha/b=-1/f$. CONCAVE.

LENSES.
Figure 10.

Surfaces at which reflections occur are also subject to this summation rule, as may be shown by using rule 2 to write down the expressions for successive refractions and reflections at a series of surfaces. It is in order that rule 3 (*b*) may be applicable that the factor n has been inserted in each of the terms of the equation for reflection at a surface, and it may perhaps be well to anticipate inquiry as to why on reflection, which involves no change of medium, it is necessary to retain n in the expression for the converging power. In this connection the following example is instructive. In figure 11 is shown a system having one

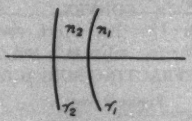


Figure 11.

refracting (r_1) and one reflecting (r_2) surface whose converging power

$$= \left\{ \frac{n_1-n_2}{r_1} + \frac{2n_2}{r_2} + \frac{-n_2+n_1}{r_1} \right\}$$

$$= \frac{2n_1}{r_1}$$

if $r_2=r_1$.

This result is independent of n_2, so that it is immaterial whether the surfaces of equal radii be separated by a parallel-sided infinitely thin non-converging layer of air, whether the reflecting surface be formed by metallic deposition on the refracting surface, or whether such a deposition be made, in the manner usual with a mirror, on the back surface of a piece of glass. In either instance the actual value of the converging power ($2n_1/r_1$) ascribed to the reflecting surface is consistent with rules 3 (*a*) and 3 (*b*).

§4. A NOTE ON RAY DIAGRAMS

Having shown by calculation that in each of the instances considered the position of the image of an object is determined uniquely by the position of the object, we may draw diagrams of the usual character illustrating the formation of images of extended objects. In these it is customary to show the path of one ray from a point on the object through the centre of curvature of a surface or the centre of a lens, together with the path of the ray which, proceeding from the same point parallel to the axis, after reflection or refraction is associated with a focal point. It should, however, be emphasized that while these two rays *locate* the image, they are not solely responsible for its formation, and to this end the paths of other rays should be added to the diagram.

Special diagrams, frequently omitted from elementary texts, require to be drawn to show the formation of the images of far distant objects by parallel rays inclined to the axis. Examples of such diagrams, the drawing of which depends upon the location of the point of intersection of an undeviated ray with the appropriate focal plane or, alternatively, with another ray which after refraction is rendered parallel to the axis, are given in figure 10.

§5. NEWTON'S FORMULA

Some teachers prefer to introduce the object-image relation in the form of Newton's equation, which may be expressed as

$$|xx'| = |f_1 f_2|,$$

where $\binom{x}{x'}$ is the distance between the $\binom{\text{object*}}{\text{image}}$ and the $\binom{\text{object}}{\text{image}}$ distant $\binom{f_1}{f_2}$ from a system that $\binom{\text{would produce}}{\text{has been produced by}}$ an $\binom{\text{image}}{\text{object}}$ at infinity. It will be more in accordance with the method of the present paper to write the equation as

$$C_1 . C_2 = -P^2,$$

where C_1 and C_2 are initial and final convergences of the light *with respect to the focal planes* distant f_1 and f_2 from a system of power P. When the light does not actually cross the appropriate focal plane or reach it without deviation the question as to its convergence or divergence with respect to this plane may be settled by supposing the rays of the beams to be produced and marked with arrows in the actual direction in which the light is travelling. Thus, for example, when light diverges from an object placed nearer to a convex lens than the focus, rays incident on the lens when produced backwards indicate that there is a converging beam passing the focal plane, and to this converging beam we may suppose the actual object to be due.

Proofs of the formula are readily obtained by the usual method, which involves the use of similar triangles: the formula may also be deduced from the diagrams of figures 1, 2 and 6 in the manner indicated below.

* The term "object" includes the image that would be formed by convergent light incident on the system supposing the system were absent.

Reflection at spherical surfaces, figure 1. In a concave mirror an object at nfinity will produce an image at a distance of half the radius from the mirror, and an object placed at the same point will produce an image at infinity. Thus in figure 1 (*b*)

$$l=f+x,\ l'=f+x',$$

and substituting in

$$1/l'=-1/l+1/f$$

shows that, in general,

$$\left(-\frac{n}{x}\right)\left(\frac{n}{x_1}\right)=-\left(\frac{n}{f}\right)^2$$

if account be taken of n, the refractive index of the material in which the reflecting surface is situated. In figure 1 (*c*) the incident light must be considered as converging, while the reflected is obviously diverging, so that

$$l=f-x,\ l'=x'-f \text{ and } \left(\frac{n}{x}\right)\left(-\frac{n}{x'}\right)=-\left(\frac{n}{f}\right)^2:$$

and in figure 1 (*d*), with the incident light converging and the reflected light diverging,

$$l=x-f,\ l'=f-x',$$

and we have

$$\left(\frac{n}{x}\right)\left(-\frac{n}{x'}\right)=-\left(\frac{n}{f}\right)^2 \text{ again.}$$

With a convex mirror an object at infinity produces a virtual image at a distance f from the mirror, and incident light converging to a point at the same distance from the mirror gives rise to a parallel beam. In figure 1 (*e*) the incident beam is diverging and the reflected beam, according to our convention, is converging: thus

$$l=x-f,\ l'=f-x',$$

and it may be shown that

$$\left(-\frac{n}{x}\right)\left(\frac{n}{x'}\right)=-\left(\frac{n}{f}\right)^2.$$

For the example shown in figure 1 (*f*) the incident beam must be taken as diverging and the reflected beam as converging as they "pass" the focal plane, and if we write $l=f-x$ and $l'=x'-f$, we have on substituting

$$\left(-\frac{n}{x}\right)\left(\frac{n}{x'}\right)=-\left(\frac{n}{f}\right)^2.$$

Refraction at spherical surfaces, figure 2. For a concave surface the power is $-(n_2-n_1)/r$, so that to produce an image at infinity (final convergence zero) the incident light must be converging to a plane whose distance f_1 from the left of the surface is given by

$$0=\frac{n_1}{f_1}-\frac{n_2-n_1}{r},$$

i.e.

$$f_1=\frac{n_1 . r}{n_2-n_1}.$$

In a similar way it follows that when the object is at infinity the light after refraction appears to be diverging from a point at a distance f_2 or $n_2 r/(n_2-n_1)$ to the right of the surface. Thus, for the example shown in figure 2 (*a*),

$$l=x-f_1,\ l'=f_1-x',$$

and on substituting in the equation there given

$$\frac{-n_2}{n_2 r/(n_2-n_1)-x'}=\frac{-n_1}{x-n_1 r/(n_2-n_1)}-(n_2-n_1)/r,$$

which on reduction gives

$$\left(-\frac{n_1}{x}\right)\left(\frac{n_2}{x'}\right)=-\left(\frac{n_2-n_1}{r}\right)^2.$$

It should be noted that the incident and refracted beams are divergent and convergent respectively when they "pass" the appropriate planes. For the example shown in figure 2 (*b*)

$$l=f_1-x,\ l'=x'-f_2$$

and we deduce

$$\left(-\frac{n_1}{x}\right)\left(\frac{n_2}{x'}\right)=-\left(\frac{n_2-n_1}{r}\right)^2$$

with the incident beam diverging (it would, in the absence of refraction, come to a focus nearer the surface than f_1) and the refracted beam converging.

For an image to be produced at infinity by a convex surface the object-distance f_1 is given by

$$0=-\frac{n_1}{f_1}+\frac{(n_2-n_1)}{r},$$

i.e. the incident light must be diverging from a point distant $n_1 r/(n_2-n_1)$ to the right of the surface. Again when the object is at infinity the light after refraction converges to a point distant f_2 or $n_2 r/(n_2-n_1)$ to the left of the surface. Thus for the example illustrated in figure 2 (*c*),

$$l=f_1-x,\ l'=f_2-x',$$

a diverging refracted beam cuts, and an incident converging beam "passes" the appropriate focal planes, and it may be shown that

$$\left(\frac{n_1}{x}\right)\left(-\frac{n_2}{x'}\right)=-\left(\frac{n_2-n_1}{r}\right)^2.$$

In figure 2 (*d*) the incident and refracted beams are converging and diverging respectively as they pass the focal planes, so that

$$l=x-f_1,\ l'=f_2-x',$$

and it may be shown that

$$\left(\frac{n_1}{x}\right)\left(-\frac{n_2}{x'}\right)=-\left(\frac{n_2-n_1}{r}\right)^2.$$

Refraction by a thin lens, figure 6. We shall suppose for simplicity that the lens is situated in air, and drop the factor n_1, § 3 (C'), from each term of our

expressions. For a concave lens an object at infinity produces a virtual image at a distance f measured to the right of the lens, whilst an incident beam converging to a point distant f to the left of the lens is by refraction rendered parallel. On the other hand, an object at infinity produces a real image at a distance f on the left-hand side of a convex lens, and an object situated at a distance f to the right of such a lens produces an image at infinity.

The conditions corresponding to the examples shown in figure 6 are tabulated in the following table.

Example	Sign of convergence of incident beams with respect to focal planes	Sign of convergence of refracted beams with respect to focal planes	Value of l	Value of l'	Equation deduced
(*a*)	−	+	$x-f$	$f-x'$	$(-1/x)\,(1/x')=-1/f^2$
(*b*)	−	+	$f-x$	$x'-f$	$(-1/x)\,(1/x')=-1/f^2$
(*c*)	+	−	$f+x$	$f+x'$	$(1/x)\,(-1/x')=-1/f^2$
(*d*)	−	+	$f+x$	$f+x'$	$(-1/x)\,(1/x')=-1/f^2$
(*e*)	+	−	$f-x$	$x'-f$	$(1/x)\,(-1/x')=-1/f^2$
(*f*)	+	−	$x-f$	$f-x'$	$(1/x)\,(-1/x')=-1/f^2$

It is thus apparent that for reflection and refraction at a spherical surface and refraction through a lens

$$C_1 . C_2 = -P^2,$$

where the symbols have the significance previously defined, and it should be noted that as P^2 is essentially positive, the left-hand side of the equation is always negative, i.e. C_1 and C_2 are of opposite sign. Thus it transpires again that for any of these systems there is a unique position of the image for any fixed position of an object, so that if P is known, C_2, and therefore the position of the image, may be found.

§ 6. TREATMENT OF COMBINATIONS OF SEPARATED LENSES OR REFRACTING SURFACES

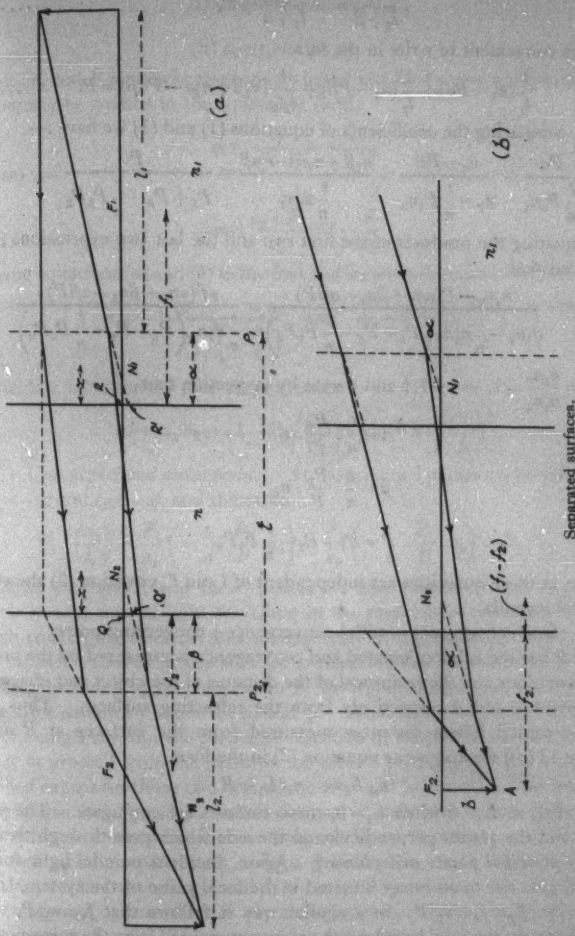
(*A*) *The location of the image of an object and of the focal points.* In figure 12 (*a*) let the medium having refractive index n be separated from the media of refractive indices n_1 and n_2 by single surfaces or by series of sensibly coincident surfaces whose effective powers are P_1 and P_2 respectively. Light from the object distant l_1 from the first surface would, after refraction, form an image at a distance l from the surface, but, on account of refraction at the second surface, we shall suppose that a real inverted image is formed at a distance l_2 from the second surface.

The equations governing the refractions are

$$n/l = -n_1/l_1 + P_1 \text{ and } n_2/l_2 = n/(l-t) + P_2,$$

and, if l be eliminated between them, it follows that

$$\frac{1}{l_1}\left(n_1 - \frac{t}{n}P_2 n_1\right) + \frac{1}{l_2}\left(n_2 - \frac{t}{n}P_1 n_2\right) + \frac{1}{l_1 l_2}\left(\frac{t}{n}n_1 n_2\right) = P_1 + P_2 - \frac{t}{n}P_1 P_2 \quad \ldots(1).$$



Separated surfaces.
Figure 12.

Now let us assume * that the quantities are related by an equation of the type

$$\frac{n_2}{l_2+\beta}=-\frac{n_1}{l_1+\alpha}+P, \qquad \ldots\ldots(2),$$

which it is convenient to write in the form

$$\frac{1}{l_1}(n_1-P\alpha)+\frac{1}{l_2}(n_2-P\beta)+\frac{1}{l_1l_2}(n_1\beta+n_2\alpha-P\alpha\beta)=P. \qquad \ldots\ldots(3).$$

Then, by comparing the coefficients of equations (1) and (3) we have :—

$$\frac{n_1-P\alpha}{n_1-\frac{t}{n}P_2n_1}=\frac{n_2-P\beta}{n_2-\frac{t}{n}P_1n_2}=\frac{n_1\beta+n_2\alpha-P\alpha\beta}{\frac{t}{n}n_1n_2}=\frac{P}{P_1+P_2-\frac{t}{n}P_1P_2}=\gamma \text{ (say)},$$

and by equating the products of the first two and the last two expressions it may be observed that

$$\gamma^2=\frac{n_1n_2-P(\alpha n_2+\beta n_1-\alpha\beta P)}{n_1n_2-\frac{t}{n}n_1n_2\left(P_1+P_2-\frac{t}{n}P_1P_2\right)}=\frac{P(\alpha n_2+\beta n_1-\alpha\beta P)}{\frac{t}{n}n_1n_2\left(P_1+P_2-\frac{t}{n}P_1P_2\right)}.$$

Thus $\gamma^2=\frac{n_1n_2}{n_1n_2}=1$, $\gamma=+1$,† and we see by inspection that

$$\alpha=\frac{t}{n}\cdot\frac{P_2}{P}\cdot n_1,$$

$$\beta=\frac{t}{n}\cdot\frac{P_1}{P}\cdot n_2,$$

$$P=P_1+P_2-\frac{t}{n}P_1P_2.$$

Moreover, as these quantities are independent of l and l', equation (2) shows that our general formula

final convergence = initial convergence + converging power

is valid if P has the value calculated and convergence is measured by the product of refractive index and the reciprocal of the distance of the object and image from points distant α and β respectively from the refracting surfaces. Thus if we denote by capital letters distances measured from the surfaces at R and Q (see figure 12 (a)) we may write equation (2) in the form

$$n_2/L_2=-n_1/L_1+P$$

and note that, as $L_2=0$ when $L_1=0$, these surfaces are conjugate. The points R and Q and the planes perpendicular to the axis which pass through them are known as *principal points and planes*.‡ Again, incident parallel light will, by refraction, give rise to an image situated in the focal plane of the system, so that when $L_1=\infty$, $L_2=f_2=n_2/P$: in a similar way it follows that $f_1=n_1/P$, where f_1 and f_2 represent the focal lengths of the system measured from the corresponding principal points.

* For a note on this assumption see Irons, *Proc. Phys. Soc.* **52**, 184 (1940).

† The negative sign is inadmissible, for when $P_2=0$, $P=P_1$.

‡ More exactly we should speak of surfaces rather than planes throughout our argument.

(*B*) *Transverse magnification.* General formulae for refraction at a surface are

(i) magnification $m=c_1/c_2$

and

(ii) $c_2=c_1+P$,

where c_1, c_2 and P represent respectively initial and final convergences and power. Thus m may be written in the equivalent forms

$$m=\frac{c_1}{c_1+P}=\frac{1}{1+\frac{P}{c_1}} \qquad \ldots\ldots(a),$$

and

$$m=\frac{c_2-P}{c_2}=1-\frac{P}{c_2} \qquad \ldots\ldots(b).$$

Applying equations (*a*) and (*b*) to the first and second refractions respectively gives

$$m_1=\frac{1}{1+P_1\Big/\left(-\frac{n_1}{l_1}\right)} \quad \text{and} \quad m_2=1-P_2\Big/\left(\frac{n_2}{l_2}\right)$$

and yields a value for the total magnification

$$M=m_1m_2=\left\{1-P_2\Big/\left(\frac{n_2}{l_2}\right)\right\}\Big/\left\{1-P_1\Big/\left(\frac{n_1}{l_1}\right)\right\}.$$

(*C*) *Unit planes and nodal points.* At the principal planes we have $L_1=L_2=0$, i.e. $l_1=-\alpha$ and $l_2=-\beta$, and therefore

$$P_1\Big/\left(\frac{n_1}{l_1}\right)=\frac{P_1}{n_1}\left(-\frac{t}{n}\cdot\frac{P_2}{P}\cdot n_1\right)=\frac{P_2}{n_2}\left(-\frac{t}{n}\cdot\frac{P_1}{P}\cdot n_2\right)=P_2\Big/\left(\frac{n_2}{l_2}\right).$$

It follows that $m=1$, and these planes are also planes of unit magnification. That is to say, rays of light travelling in the medium of refractive index n_1 as though converging towards, and subsequently diverging from, a point distant y (the "object") above the axis at R are so refracted that on traversing the medium of refractive index n_2 they appear to be diverging from a point (the "image") at the same distance y above the axis at Q. What is true of a series of rays is also true of a single ray, but the particular direction in which any one incident ray will appear to proceed from the point on the second principal plane cannot as yet be specified except for those two rays which after and before refraction pass through the principal foci F_2 and F_1, and at one part of their path are parallel to the axis. Now although these two rays are sufficient to locate the image of an object formed by either a compound system or by a thin lens (in which the principal planes coincide as t, and therefore α and β, are zero) it is instructive to inquire whether in a compound system there is any analogue to the optical centre of a thin lens as this point, having the property that light passes through it undeviated, is commonly employed in simple ray-constructions. To this end we shall postulate two points N_1 and N_2 each situated on the axis at a distance x from the corresponding principal plane as shown in figure 12 (*a*), and suppose that rays travelling

in the medium of refractive index n_1 towards N_1 after refraction by the system appear to diverge from N_2. The value of x may then be determined from the standard equation

$$-\frac{n_2}{x}=-\frac{n_1}{x}+P,$$

which indicates that

$$x=\frac{n_1}{P}-\frac{n_2}{P}=f_1-f_2$$

and that the points N_1 and N_2 coincide with the principal points when $n_1=n_2$.* Moreover, as $QQ'=RR'$ as well as $QN_2=RN_1$ (see the figure), an incident ray directed towards N_1 and the corresponding refracted ray directed from N_2 must be parallel, and use may be made of the points N_1 and N_2, which are known as nodal points, in locating the position of an image. It should perhaps be added that these points must lie on the axis of the system, since otherwise a ray directed towards N_1 and parallel to the axis would emerge parallel to the axis and not pass through F_2. Once the image has been located it is possible, as with a thin lens, to draw the path of a ray proceeding in any direction from a point on the object to the corresponding point on the image.

(*D*) *Power.* It remains to show that the quantity P, which we have shown to be equal to $P_1+P_2-(t/n)P_1P_2$, n_1/f_1 and n_2/f_2 and have called the "power," has the significance previously attached to this term. This may be done by considering the paths through the system of two rays supposed to emanate from a point on a distant object (see figure 12 (*b*)). One of these rays passes through the nodal point N_1 and emerges parallel to its original direction through N_2, cutting the focal plane at A, while the other, like all similar rays, also passes through A, as the image is formed in the focal plane. By definition the power is $\alpha n_1/(AF_2)$ which, as $\alpha=(AF_2)/(f_2+f_1-f_2)$, is the same as n_1/f_1 and therefore indicates the consistency of the nomenclature.

The formulae for α and β and for the magnification have been obtained on the supposition that the system acts in the manner shown in figure 12 (*a*), and, that in the expressions for the initial and final convergences, α and β have to be added to the distances of the object and its image from the surfaces. The formulae are, however, quite general, and the interpretations to be placed on negative values of P, α and β are explained in the solutions of problems (8) and (9) in the next section.

Newton's equation holds equally for a lens system and a single thin lens, and the proof follows the lines already described for the example illustrated in figure 6 (*d*).

* If in figure 12 (*a*) the points were on the left of the principal planes, the equation would be

$$\frac{n_2}{x}=\frac{n_1}{x}+P$$

with $$x=\frac{n_2}{P}-\frac{n_1}{P}=f_2-f_1:$$

the points are therefore displaced towards the side on which the end medium is the denser.

§7. ILLUSTRATIVE EXAMPLES

The following worked examples are inserted merely to illustrate the general use of the methods proposed in the paper, and no claim is made that the solution of any one problem is the most appropriate or instructive that can be devised.

The symbols c_1, c_2, P and m will be used to denote initial and final convergence, power and magnification respectively.

(1) *Find the distance between the images of his eye observed by a man looking into a sheet of glass which has its back surface silvered* (figure 13 (*a*)).

(*a*) The image formed by reflection at the front surface is as far behind the surface as the object is in front, i.e. $l_1=l$.

(*b*) As the powers of the surfaces are both zero, the light is always diverging, so that for

(i) the first refraction

$$-\frac{n}{l_2}=-\frac{1}{l}+0, \quad \therefore \quad l_2=nl.$$

(ii) reflection at the back surface

$$-\frac{n}{l_2'}=-\frac{n}{l_2+t}+0, \quad \therefore \quad l_2'=l_2+t.$$

(iii) the second refraction

$$-\frac{1}{l_2''}=-\frac{n}{l_2'+t}+0, \quad \therefore \quad l_2''=\frac{l_2'+t}{n}.$$

$\therefore$ Required distance

$$=l_2''-l_1=\frac{nl+2t}{n}-l=\frac{2t}{n}.$$

(2) *It is required to produce an image three times the length of an object by means of a mirror of power 4D. Show that there are two possible positions of the object and find the distance between them.*

The positions correspond to $m=\pm 3$.

(*a*) Erect virtual image

$$m=c_1/c_2=+3; \quad c_2=c_1+4; \quad \therefore \quad c_1/3=c_1+4; \quad c_1=-6D,$$

i.e. light diverges from a point $100/6=16\frac{2}{3}$ cm. from mirror.

(*b*) Inverted real image

$$m=c_1/c_2=-3; \quad c_2=c_1+4; \quad \therefore \quad -c_1/3=c_1+4; \quad c_1=-3D,$$

i.e. light diverges $100/3=33\frac{1}{3}$ cm. from mirror,

$\therefore$ Distance between objects $=16\frac{2}{3}$ cm.

(3) *What is the power of an auxiliary lens that must be placed close to an eye from which the crystalline lens has been removed in order that the patient may view distant objects?* (*The cornea may be considered as having a radius of 7 mm. and to be separated by a uniform medium of refractive index 4/3 from the retina distant 22 mm. from it.*)

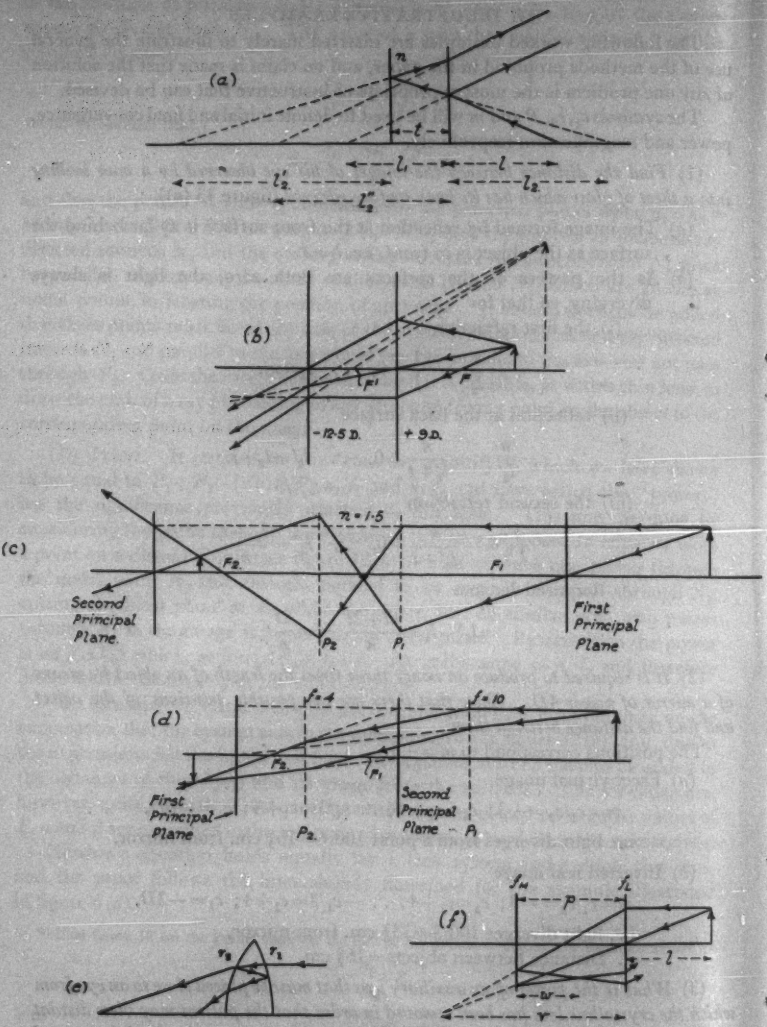

(a)
n
t
l1
l
l2
l'2
l''2
(b)
F'
F
-12·5 D.
+ 9 D.
(c)
n = 1·5
F2
F1
Second Principal Plane.
First Principal Plane
P2
P1
(d)
f = 4
f = 10
F2
F1
First Principal Plane
Second Principal Plane
P2
P1
(e)
r2
r1
(f)
fH
fL
p
l
Δr
Y

Figure 13.

Let the power of the lens required be P, so that the total power of the lens and convex refracting surface is $P+(-1+4/3)/0{\cdot}007$. Applying our general equation we find that $(4/3)/(0{\cdot}022)=0+(P+1/0{\cdot}021)$, i.e. $P=+13D$. (Having regard to the curvature of the cornea the result is necessarily approximate.)

(4) *What spectacles should be prescribed for a person who wishes to view distant objects but can only see distinctly objects which are between* 12·5 *and* 25 *cm. from his eye; and what will be his least distance of distinct vision when he is wearing these spectacles?*

He must be provided with a lens such that rays from a distant object ($c_1=0$) after refraction must appear to diverge from a point 25 cm. from his eye ($c_2=-4D$), i.e. a $-4D$ lens.

If his new least distance of distinct vision is d_n from his eye, light whose initial convergence is $-1/d_n$ must after refraction appear to diverge from 12·5 cm. from his eye, i.e. $-8=-1/d_n-4$ or $d_n=\frac{1}{4}$ metre$=25$ cm.

(Alternatively we may equate values of accommodating power, i.e. put $100/12{\cdot}5-100/25=1/d_n-1/\infty$ or $d_n=\frac{1}{4}$ metre.)

(5) *A right-angled triangle is placed with its side of length* 4 *mm. along, and its side of length* 3 *mm. perpendicular to and above the axis of a lens which is* 50 *cm. distant. The lens is made of glass of refractive index* 1·5 *and both its surfaces are concave towards the triangle, the nearer having a radius of* 100 *cm. and the more remote a radius of* 10 *cm. Describe the image of the triangle formed by the lens.*

As the power of the lens is $(+1-1{\cdot}5)/1+(1{\cdot}5-1)/0{\cdot}1=4{\cdot}5\ D$, $c_2=-2+4{\cdot}5=2{\cdot}5$ and a real image is formed 40 cm. on the opposite side of the lens to the object. The transverse magnification $=c_1/c_2=-2/2{\cdot}5=-4/5$, i.e. the image of the triangle is below the axis and the image of the side of length 3 mm. is 2·4 mm. The longitudinal magnification $=1.(c_1/c_2)^2=16/25$, i.e. the image of the side of length 4 mm. is 2·6 mm. long, and as c_1/c_2 is negative the part of the triangle nearer the lens in the object is more remote in the image.

(6) *A, B and C are three points in a straight line such that* $AB=20$ *cm. and* $BC=15$ *cm. If an object of length* 1 *cm. and lenses of powers* $+9D$ *and* $-12{\cdot}5D$ *are placed perpendicular to the line at A, B and C respectively, find the position, the size and the nature of the image formed by the system* (figure 13 (*b*)).

For refraction at B, $c_2=-5+9=4D$, $m=c_1/c_2=-5/4$, i.e. in the absence of the lens at C an inverted real image of length 5/4 cm. would be formed 25 cm. from B. For refraction at C, $c_2=+10-12{\cdot}5=-2{\cdot}5\,D$, $m=+10/-2{\cdot}5=-4$, i.e. after refraction at C the light appears to diverge from a virtual image 40 cm. away. The total magnification is $+5$, so that the image is 5 cm. long and erect with respect to the object.

(7) *Light from a distant object is incident on a thin double convex lens and forms images on the opposite side of the lens at distances of* 70 *and* 10 *cm. from it. If the first of these is formed in the usual way and the second after two internal reflections, calculate the refractive index of the material of the lens* (figure 13 (*e*)).

Power of lens for light travelling directly through it,

$$P_1=(-1+n)/r_1+(n-1)/r_2.$$

Power of lens for light internally reflected,

$$P_2=(-1+n)/r_1+2n/r_2+2n/r_1+(n-1)/r_2.$$

$P_1/P_2=(n-1)/(3n-1)=$(final convergence)$_1\div$(final convergence)$_2=10/70$; whence $n=1{\cdot}5$.

*(8) *Find the positions of the unit and focal points for a glass lens of refractive index* 1·5 *and thickness* 7 *cm., the radii of the convex refracting surfaces being each equal to* 1 *cm.*

Keeping lengths in cm. and substituting in the equations given in §§ 6 (*A*) and 6 (*C*) gives

$$P=P_1+P_2-(t/n)P_1P_2=(-1+1{\cdot}5)/1+(1{\cdot}5-1)/1-(7/1{\cdot}5)(\tfrac{1}{2})(\tfrac{1}{2})=-\tfrac{1}{6},$$

$$\therefore\ |f|=6 \text{ cm.},$$

$$\alpha(=\beta)=(t/n)(P_2/P)n_1=(7/1{\cdot}5)(\tfrac{1}{2})(-6)1=-14,$$

and the nodal and principal points coincide.

Now these equations have been deduced for a converging system with the distance α *added* to the distance of the object from the surface to obtain the initial convergence; our results therefore indicate that in this arrangement the first principal plane is situated 14 cm. nearer a distant object than the first refracting surface; that the second principal plane is similarly outside the second refracting surface and, further, that the system is a diverging one. This means that a ray originally parallel to the axis after refraction will diverge from a point on the second principal plane away from the axis, so that F_2 is nearer the lens surface than the principal plane by 6 cm.: F_1 is similarly situated with respect to the first principal plane. The formation of an image by the system is illustrated in figure 13 (*c*).

(9) *An eyepiece is made of two thin convergent lenses of focal lengths* 4 *cm. and* 10 *cm. respectively at a distance of* 7 *cm. apart. Calculate the focal length of the eyepiece, and draw a diagram to indicate the position with regard to the lenses of the focal and unit points of the eyepiece.*

Applying our formulae as in the previous example and working with lengths in metres, we have

$$P=10+25-0{\cdot}07\,.\,10\,.\,25=17{\cdot}5\ D.\quad f=100/17{\cdot}5=5{\cdot}71 \text{ cm.}$$

$$\alpha=0{\cdot}07\ (25/17{\cdot}5)=+0{\cdot}1 \text{ metres}=10 \text{ cm.}$$

$$\beta=0{\cdot}07\ (10/17{\cdot}5)=+0{\cdot}04 \text{ metres}=4 \text{ cm.}$$

As the lengths involved in the expressions for the initial and final convergences are obtained by *adding* 0·1 and 0·04 metres to the distances of the object and its image from the first and second lenses, the situations of the principal planes

* Problems 8 and 9 are taken from the *Report on the Teaching of Geometrical Optics* (published by the Physical Society), which contains a number of alternative solutions.

are as indicated in figure 13 (*d*), which also shows the formation of an image by the system. Again, as the system is converging, a ray travelling from right to left and originally parallel to the axis on leaving the second principal plane passes through F_2, whose position may therefore be fixed: the position of F_1 may likewise be located by considering a similar ray travelling from left to right.

It is an instructive drawing exercise to locate the position of the image of an object formed by such a compound system, (*a*) by making use of the properties of principal planes and nodal points whose positions are calculable from our general formulae, and (*b*) by finding in order the positions of the images formed by the first and second lenses.

(10) *As a final example we shall consider a system consisting of a thin lens placed in front of a mirror, and determine its self-conjugate point* (figure 13 (*f*)).

For the first refraction through the lens, suppose that a real image is formed so that

$$1/v=-1/l+1/f_L \qquad \ldots\ldots(1),$$

where v is the distance of the image from the lens and, if we suppose that this is greater than p,* the distance between the lens and mirror, we have for reflection at the mirror

$$\pm 1/w=1/(v-p)+1/f_M \qquad \ldots\ldots(2),$$

according as the image is formed at a distance w in front or behind the mirror. Finally, for refraction through the lens again we have, as the object and image are to coincide,

$$1/l=-1/(p\mp w)\dagger+1/f_L \qquad \ldots\ldots(3),$$

according as the mirror produces a converging or diverging beam. From (1) and (3) we note that

$$v=p\mp w. \qquad \ldots\ldots(4),$$

The total magnification produced will be given by

$$m=\left(-\frac{1}{l}\Big/\frac{1}{v}\right)\left(\frac{1}{v-p}\Big/\pm\frac{1}{w}\right)\left(-\frac{1}{p\mp w}\Big/\frac{1}{l}\right)=\left(\frac{v}{v-p}\right)\left(\frac{\pm w}{p\mp w}\right)=|1| \qquad \ldots\ldots(5).$$

By appropriate substitutions in these equations we may deduce the theory of some well-known methods for finding the constants of lenses and mirrors. Suppose that :—

(*a*) The image formed by the lens is at infinity, i.e. there is no image other than the final one.

$v=\infty$; $\therefore$ by (4) $w=\infty$, by (1) $f_L=l$, by (2) $f_M=\infty$, and by (5) $m=-1$, i.e. the image is inverted. (These conditions correspond to the ordinary method of finding the focal length of a convex lens by means of a plane mirror.)

* If $p>v$, the initial convergence at the mirror is $-1/(p-v)$, and is unaltered.

† If $w>p$ and a real image were formed by the mirror, converging light would fall on the lens, the initial convergence would be $+1/(w-p)$, and equation (3) would still satisfy the conditions.

(*b*) The image formed by the lens is behind the mirror, i.e. $v>p$.

Thus in equation (4) the + sign must be taken; therefore the mirror produces a diverging beam, and $v=p+w$. Substitution of this in equation (2) gives $f_M=-w/2$, i.e. the mirror must be convex and the image must be formed at its centre of curvature. Moreover, $m=-1$ by equation (5), so that the image is inverted. (Method for finding the radius of a convex mirror by means of a convex lens.)

(*c*) The image is formed by the lens on the mirror; or $v=p$.

Substitution in equation (1) gives $1/p=-1/l+1/f_L$ and in equation (4) gives $w=0$. Putting $w=0$ and $v=p$ in equations (2) and (5) shows that f_M and m are indeterminate. If, however, the image is formed by the lens a very small distance h in front of or behind the mirror, the right-hand side of equation (2) may be modified so that $\pm 1/w=\mp 1/h+1/f_M$ and, unless f_M is very small, the terms of the equation not involving f_M being large, must be equal in sign and magnitude so that the magnification produced by the reflection is +1. Thus

$$m=\left(-\frac{1}{l}\Big/\frac{1}{p}\right).1.\left(-\frac{1}{p}\Big/\frac{1}{l}\right)=+1,$$

and retains this value as $\pm h\to 0$, indicating that the final image is erect. (This analysis explains a method for determining the focal length of a convex lens which deserves greater popularity than it appears to possess, on account of the accuracy with which settings may be made. In optical-bench tests with a luminous source, however, it is usually necessary to place a strip of glass just in front of the source, and at 45° to the axis of the system so as to bring the reflected light to a focus on a ground-glass screen at the same distance away from the reflector as the source. It may also be noted that as f_M is indeterminate the curvature of the mirror employed to reflect the light back to the source is immaterial. This result is of interest in connection with accounts of Fizeau's experiment for the determination of the velocity of light, since the reflector at the distant station is as often as not stated to have been concave.)

(*d*) The image is formed by the lens in front of the mirror, i.e. $v<p$.

Thus in equation (4) the negative sign must be taken, so that the mirror produces a converging beam and $v=p-w$. Substitution of this in equation (2) gives $f_M=w/2$, i.e. the mirror must be concave and form an image at its centre of curvature. Moreover, $m=-1$ by equation (5), so that the final image is inverted. Here then the two intermediate images coincide.

We might investigate likewise the properties of a system in which a virtual image is formed by the first refraction, but it is hoped that the above examples afford sufficient illustration of the method.

§8. CONCLUSION

A paper on such a well-discussed topic as simple optical calculations must necessarily contain much that its writer has culled consciously or unconsciously from the perusal of articles and textbooks. Indeed, it is on account of the

difficulty in tracing some ideas to their sources that references have been omitted to such well-known treatments as that upon which §3 (C') is based. It is nevertheless hoped that some of the contents, for example the simple proofs of the standard formula which involves finite objects, the insistence on the truth of this formula for both divergent and convergent light, the treatment of Newton's formula, and the discussion of lens systems, may be sufficiently novel for the paper to serve the dual purpose of stimulating interest in a non-Cartesian method and providing a treatment that may be extended or simplified as desired.

POLISH LAYERS ON NICKEL

By P. E. AXON, B.Sc., A.R.C.Sc., D.I.C.

Communicated by Prof. G. P. Thomson, F.R.S., 23 October 1939

ABSTRACT. Surface layers which give an electron-diffraction pattern of diffuse haloes have been produced on nickel by several methods. The methods differ in the amount of abrasion they employ, and in the pressure applied perpendicular to the nickel surface as abrasion is carried on. Surfaces are created by abrasion with emery, in one case by hand rubbing and in another by spinning the specimen in a lathe chuck. Yet another method employs a cylindrical nickel bullet of carefully chosen dimensions which is etched and fired from a revolver into a tank of water. The tips of the rifling ridges of the revolver barrel lie on a circle of slightly less diameter than the circular cross-section of the bullet. When the bullet is fired, the rifling ridges create areas which are, in effect, subjected to a single polishing stroke at very great abrasive pressure. Although, by virtue of the cross-sectional diameter of the bullet, the surface areas between the rifling ridges are not subjected to abrasion, a layer is created which gives an electron-diffraction pattern of diffuse haloes. These diffuse haloes probably result from a smoothing effect having been carried out by melting of the projections, due to heat transferred from the gaseous products of explosion. The depth and chemical composition of this and the other layers are found by removing them by controlled etching and sputtering in an argon discharge chamber. Results show that abrasive action has more extensive effects than mere surface smoothing. The depth of amorphous layers created by the abrasive methods appears to be a function of (1) the amount of abrasion, and (2) the pressure perpendicular to the surface. The influence of the conditions of production on the depth and chemical composition is discussed. Increase of abrasive pressure seems to increase the oxide content in the layers, and this may be due to adsorbed oxygen layers on the metal surfaces and to higher temperatures. The Bragg spacings of the diffuse haloes obtained from the various surfaces are compared with results obtained by Dobinski, Plessing and others. It seems that nickel, under polishing conditions not yet clear, can give haloes with Bragg spacings which correspond to one or other of two approximate but distinct sets of values.

§1. INTRODUCTION

The structure of polish on metals has been a controversial subject. Hitherto most of the polish layers investigated have been produced by the same means, namely, by abrasion with emery and chamois leather or with rouge and cloth pads. The electron-diffraction pattern from the polished layer so produced consists of two diffuse haloes. This pattern was interpreted by French[1] as being due to an amorphous structure in the layer. Others hold the view that the layer must consist of very small crystals. Another possibility was created by Kirchner[8] when he demonstrated that a crystalline film with few surface projections could give the diffuse halo-pattern by reflexion, whilst giving

the normal sharp ring-pattern by transmission. The latter property showed that the crystals were not small. The possible conclusion from this experiment that all polish layers gave a diffuse ring-pattern merely because their surface projections had been removed in the rubbing process was nullified by Cochrane[3]. He was able to strip the polish layer from a gold specimen and to show that its transmission pattern also consisted of two diffuse haloes. The polish layer was therefore of different texture to the very smooth surface film obtained by Kirchner.

It was thought that the more violent application of factors present in ordinary polishing processes might provide further data on the structure of polish and the mechanism producing it. With this in view, Professor G. P. Thomson suggested that the passage of a bullet along the rifled barrel of a gun might provide more violent forms of polishing action. The rubbing of a metal bullet on the metal rifling ridges would possibly exaggerate the effects of abrasion on the metal surfaces. Experiments arising from this, and the effects found, are here described.

§ 2. EXPERIMENTS

The gun employed for the experiments described was a Colt Automatic, calibre 0·32. The first experiments were carried out with the ordinary variety of bullet issued by manufacturers for this calibre.

The bullet was extracted from its cartridge case and etched in 10-per-cent ammonium persulphate for a suitable time determined by trial. The electron-diffraction pattern from the casing of the bullet consisted of rings which could be attributed to nickel and chromium. The casing apparently consists of some alloy of these two metals. The bullet after removal from the camera was again inserted into the cartridge case without any of the parts later to be examined being touched with metal tools. After soaking in benzene to remove grease, the bullet was loaded into the gun and fired into a tank of water about five feet deep. It was retrieved from the bottom, soaked in alcohol, and examined in the electron-diffraction camera. The bullet gave patterns as follows: (1) near the nose an almost unchanged sharp ring-pattern, and (2) near the edge, i.e. parts near the barrel, definite signs of two diffuse rings.

The experiment is difficult to interpret. The diffuse rings near the rim may have been due to rubbing on the rifling, to hot gases generated in the explosion, or to both. It was of value, however, in demonstrating that polishing effects were present. A new shape of bullet was then designed and made in the workshops. On this it was more easy to separate out the various effects. The new form was a pure nickel cylinder which could be turned on a lathe.

In figure 1, A is a small cylindrical projection, B and C are of equal diameter, and a groove is sunk, as shown, between B and C. The object of the projecting cylinder A was twofold: (1) it was the part which struck the bottom of the tank, and thus relieved part B from strain or violent contact with the tank; (2) it was afterwards gripped in the chuck of a lathe, and the part C was machined away until it was equal in diameter to A, and the bullet was of the form shown in figure 2.

The preparation of the bullet before firing took place when it was in the form 1. The bullet was held in a lathe chuck by the part A, and the chuck was set revolving. The part B was then ground up on (0—0000) emery, both emery paper and bullet being constantly soaked in benzene poured on from above. The emery paper was moved regularly to and fro along the axis of the bullet so that a uniform treatment of the surface resulted. The process was completed by using clean chamois leather in a similar manner. This process took about twenty minutes. At this stage the surface of part B, when examined by electron diffraction, gave a pattern consisting of two diffuse haloes, the Bragg spacings being 1·08 A. and 1·86 A. The bullet was then etched in 10-per-cent ammonium persulphate for a definite time until, on being placed in the diffraction camera once again, it gave a pattern of moderately sharp nickel rings showing signs of orientation.

The bullet was now inserted in the explosive-filled cartridge case. In order to do this the cartridge case was held in a lathe chuck and the part C of the bullet was eased in by tapping lightly on A. The end of the cartridge case was then spun into the groove between B and C, removed from the chuck, and held vertically in a clamp so that part B was in benzene. A soaking of suitable duration removed

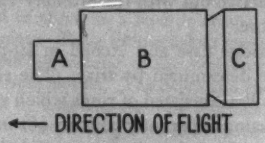


Figure 1. Preparation of specimen.

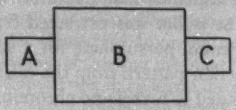


Figure 2. Finished specimen.

dust and grease which may have accumulated. The barrel of the pistol was also soaked in benzene to clean it. It is important to note that the part B of the bullet had not been touched except gently by the fingers. The bullet was now ready for firing, and was loaded into the pistol. The parts B and C, figure 1, were of equal diameter. This diameter was such that when the bullet was travelling down the barrel its periphery would lie between the base and top of each rifling ridge. Thus when the bullet was fired the ridges pressed out five distinct channels, whilst parts lying between the five ridges were not in contact with any part of the barrel. These two distinct types of bullet surface will hereafter be referred to as the "contact" and "non-contact" parts. The abrasive action in the contact parts was extremely severe, and only forces of the magnitude of those generated in an explosion could have forced the bullet down the barrel. After being fired, the bullet was retrieved from the water tank and plunged into alcohol. This removed water and soaked off powder which in some cases had been deposited on the non-contact parts. The part A, figure 1, was then held in a lathe chuck, and the part C was machined away until the bullet was of the form shown in figure 2. Part B was not touched in this process. The bullet was then again soaked in benzene before examination in the diffraction camera.

The patterns from both the contact and the non-contact parts of the bullets at this stage consisted of two diffuse haloes. The Bragg spacings were as follows:

	First halo (A.)	*Second halo* (A.)
Contact surface	$1{\cdot}83 \pm 0{\cdot}04$	$1{\cdot}11 \pm 0{\cdot}02$
Non-contact surface	$1{\cdot}87 \pm 0{\cdot}03$	$1{\cdot}12 \pm 0{\cdot}02$

A nickel block was also polished on emery paper under benzene. It was abraded on (0—0000) emery paper and finally on chamois leather, the whole process taking about twenty-five minutes. After this process an electron-diffraction pattern of two diffuse haloes was obtained from the surface. The Bragg spacings were $1{\cdot}8 \pm 0{\cdot}04$ A., $1{\cdot}0 \pm 0{\cdot}04$ A. The Bragg spacings of the two patterns obtained by polishing under benzene were thus:

	First halo (A.)	*Second halo* (A.)
Block polish	$1{\cdot}8 \pm 0{\cdot}04$	$1{\cdot}0 \pm 0{\cdot}04$
Lathe polish (as described)	$1{\cdot}86 \pm 0{\cdot}04$	$1{\cdot}08 \pm 0{\cdot}04$

Experiments were now carried out to determine the nature and depths of the various layers. This was done in two ways: (1) by controlled etching in 10-per-cent ammonium persulphate, and (2) by sputtering off the surfaces in an argon discharge chamber.

Controlled etching. The method here was to plunge the specimen into 10-per-cent ammonium persulphate for known short periods and to examine by electron diffraction the surface layer revealed by each such period. An etching process would not take the surface off quite uniformly but would tend to remove the material along crystal faces. Over a suitable area, however, this process should lead to a comparative estimate of the depths of the various layers examined. The process was continued on each surface until the diffuse halo pattern gave way to a return to sharp rings.

The orientation observable after 200 seconds in the lathe-polish pattern was not similar for all specimens. In one $\sqrt{3}$ was found lying parallel to the surface, in another $\sqrt{4}$, whilst in a third both $\sqrt{3}$ and $\sqrt{4}$ were preferentially oriented but not parallel to the surface.

These values of etching time will lead to a value of actual depth of layer. The values are tabulated later.

The sputtering process. In order to estimate the depth of the layers by this method the specimen in its form shown in figure 2 was made the cathode in an argon discharge chamber. The bombardment of the argon ions on the surface of the specimen removed material. Under constant conditions it then gave a direct measure of the depth of material removed.

In order that all parts of the cylindrical surface of the specimen should be uniformly treated, special design of the electrodes was necessary.

The cathode was made of aluminium and was of the form shown in figure 3. The specimen in its final form, figure 2, could be held by its cylindrical projections

Table 1. Surfaces removed by etching

Time (sec.)	Non-contact layer	Contact layer	Lathe-polish layer	Block-polish layer
0	Diffuse haloes	Diffuse haloes	Diffuse haloes	Diffuse haloes
5	No change	No change	No change	No change
8	No change	No change	No change	Haloes less diffuse
13	No change	No change	No change	The nickel pattern appeared. Rings broad but all resolved
20	Haloes less diffuse	No change	No change	Sharp rings
25	Sharp rings appeared. Low intensity. Strong rings of final pattern more obvious	No change	No change	
30	Sharp pattern of nickel with, in some cases, nickel oxide	No change	No change	
40		Sharp ring pattern begins to appear. The rings are broad and there is much background	No change	
45		The pattern sharpens up; weaker rings begin to appear	No change	
50 to 55		Pattern complete. Pattern consists of nickel and nickel oxide rings	No change	
120			Rings begin to appear. Broad and faint	
200			Nickel pattern. Orientation observable	

in the cradles A and B of the cathode. The upper cradle B was screwed down on to the upper part of the specimen. All parts of the cathode except the cradles and specimen itself were covered with glass so that the discharge was concentrated on to the specimen.

The anode was a short circular cylinder which, when in position, surrounded the specimen, which was cylindrical, the axes of the anode and the specimen being coincident. This symmetrical arrangement of electrodes, which is shown in figure 3, concentrated the maximum bombardment of positive ions on to the bullet surface. The outer diameters of the cradles A and B were made equal to the diameter of the specimen, so that possible edge effects would be reduced during sputtering. The electrodes were enclosed in an air-tight chamber of glass which was attached to a high-vacuum pump. When the chamber had been evacuated, argon was fed in from a cylinder through a needle valve. The argon feed was adjusted always to keep the current passing at a constant value for all surfaces and sputtering periods. The mode of procedure was to examine the

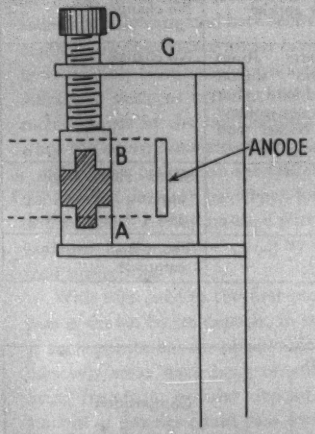


Figure 3. Elevation of electrodes and specimen in position.

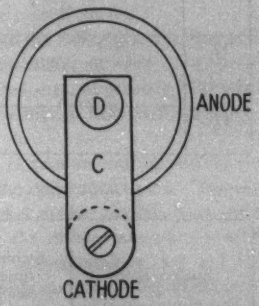


Figure 4. Plan of electrodes in position.

surface condition by electron-diffraction sputter for a known period at the constant current, and then to examine the fresh condition of the surface by electron diffraction. The time of sputtering was thus a direct measure of the depth of surface layer removed. The three surfaces on the cylindrical bullet were investigated by this means, and the general variation is given in table 2.

A calibration experiment was carried out on the sputtering apparatus so that actual values of layer-depths were calculable. A cylindrical bullet was cleaned, weighed on a microbalance, sputtered for a long known period, and then reweighed. A knowledge of the loss in weight and area of cylindrical surface gave a value of material sputtered off in unit time. The depths obtained by both etching and sputtering processes are tabulated. The value of the lathe-polish

depth, as obtained from the sputtering calibration, has been assumed for the first table, and the other depths have been obtained by direct comparison of etching times.

Table 2. Surfaces removed by sputtering

Time (min.)	Non-contact layer	Contact layer	Lathe-polish layer
½	Two diffuse haloes	Two diffuse haloes	Two diffuse haloes
1	No change	No change	No change
1½	Moderately sharp rings of nickel and nickel oxide	No change	No change
2½		Sharp rings appear faintly	No change
3		Strong pattern. Rings broad and similar to those found after 1½ min. on non-contact layer. Nickel and nickel oxide	No change
12			Sharp rings begin to appear. Broad, $\sqrt{3}$ and $\sqrt{4}$ not resolved
15			The nickel pattern complete

Table 3

Etching process

Surface layer	Etching time (sec.)	Depth (A.)	Composition
Non-contact	25 to 30	70	Nickel and nickel oxide or nickel
Contact	50 to 55	150	Nickel and nickel oxide
Lathe polish	200	600	Nickel
Block polish	13	40	Nickel

Sputtering process

Surface layer	Sputtering time (min.)	Depth (A.)	Composition
Non-contact	1½	60	Nickel and nickel oxide
Contact	2½–3	120	Nickel and nickel oxide
Lathe polish	15	600	Nickel

Hopkins [4] found a depth of amorphous layer of 40 A. for the hand-polish layer on gold. This agrees well with the depth of the hand-polish layer on the nickel block found here.

§3. CONSIDERATION OF RESULTS

The electro-diffraction pattern of two diffuse haloes obtained from polished metal surfaces has been interpreted in two ways: (1) as being due to an amorphous layer produced by the rubbing action, and (2) as being due to the removal of projections, resulting in a smooth surface.

The experiments described in this work seem to indicate that ordinary polishing has more far-reaching and extensive effects than that of merely smoothing the surface. In connexion with the first interpretation it appears that little important distinction can be made between a composition of small crystals, molecules packed as in a liquid, and a mixture of both these forms. From the point of view of electron-diffraction experiments the pattern obtained from any one of the three structures would consist of a pattern of diffuse haloes. Indeed, some layers investigated here seem to embrace all three structures according to the depth below the surface under investigation. This possibility has been advanced, however, only from considerations of the method of production of the layers. The diffraction patterns taken at different depths of the layer could not distinguish such changes in the composition. The first visible change in patterns occurs when the amorphous structure gives way to larger crystal grains.

The layer created on the bullet surface between the rifling grooves, that is the non-contact part, can have been formed in one of two different ways: (1) a removal of projections by the passage past them, at high velocity, of gaseous products of explosion and powder; or (2) by transference of sufficient heat from the gases to melt projections.

With reference to the first possibility, that a certain amount of powder does pass is shown by its deposit, in some cases, on the non-contact part. The layer at such points has an observable content of nickel oxide. As has been stated, however, areas have been obtained which give the diffuse halo-pattern whilst being free from powder deposit and oxide. It seems unlikely that any large amount of powder could pass without an observable quantity being deposited on the surfaces. We may, therefore, consider that the passage of high-velocity powder particles is not a necessary condition for the production of the surface change. The removal of this layer by the bombardment of heavy argon ions with energies of several thousand volts requires about ninety seconds. It therefore seems unlikely that the bombardment of gas molecules of thermal energy in the very short time available could remove projections.

We are thus left with the second explanation, that the projections have melted and flowed into the valleys. This would produce a smooth surface with few projections and would be expected to give a pattern of diffuse haloes in accordance with the findings of Kirchner[2]. It has been found that the depth of this layer is 60 to 70 A. This would mean that the initial crystalline projections were 120 to 140 A. high, which is a value of the order usually accepted for the projections on etched metal surfaces.

If the attainment of the polish condition were always a question of smoothing

there should be little difference between the block-polish layer and those attained by the more violent methods of abrasion. In either case the surface would once more give sharp ring-patterns when the upper material had been removed and crystalline projections were exposed. Sharp rings would appear when the projections were large enough to resolve the various rings.

The mechanism, whatever it may be, that is assigned to the production of the non-contact layer in the gun barrel cannot alter the view that this layer can be considered as a zero mark for the groove or contact layer. Thus any increase in depth of the contact layer as compared with that of the non-contact layer, or indeed that of the block-polish layer, must be directly due to an amorphous layer. In these experiments, three layers formed by abrasion were investigated: (1) the layer formed in the grooves on the bullet by great pressure acting for a short time, i.e. in a single polishing stroke; (2) the layer formed by polishing the cylindrical specimen by revolving it (this was carried on for about twenty-five minutes at six hundred revolutions per minute, i.e. about fifteen thousand polishing strokes at finger pressure); (3) the layer formed on the nickel block by polishing on a flat plate covered with emery. This was carried on for about twenty-five minutes at a stroke per second, i.e. about fifteen hundred strokes at finger pressure.

The depths of the three layers were 120 to 150 A., 600 A. and 40 A. respectively. Thus it would seem that the depth of the polish layer formed by abrasion is a function of: (1) the time of polishing, or more precisely, the number of polishing strokes, and (2) the abrasive pressure, i.e. the pressure perpendicular to the surface.

Thus the single stroke at very great pressure is able to produce a layer deeper than that due to a larger number of strokes at finger pressure, but only a fifth of the depth of that due to the very large number of strokes of the lathe polish. There is probably a limiting depth of the polish layer for any particular abrasive pressure provided that polishing is carried on for a sufficiently long time.

We can imagine that the first crystals broken by the polishing stroke act as a grinding medium for those beneath. The process of crystal-breaking would then continue until a certain limiting depth was reached, beyond which the abrasive pressure would be insufficient to cause any further fracturing.

Bowden [5] has demonstrated that temperatures in the rubbing process are high enough to cause local melting. Melted material at the top of the layer would then mix with small broken crystals. Such a mechanism would produce a layer in which there was transition from an amorphous mixture through broken crystals to normal grain-size. This would provide the first really distinct change in the electron-diffraction pattern, namely, the reappearance of sharp rings.

It would be at this point, the bottom of the amorphous layer, that orientation effects should become observable. The abrasive force, whilst not large enough to break crystals up at this depth, would still expend itself in causing crystals to adopt a position of easy-slip planes. The unfailing occurrence of nickel oxide in the polish layer formed by the rifling of the gun is also of interest. The existence

of adsorbed layers of gases on the surfaces of metals is well established. There is thus probably an adsorbed layer of oxygen on the surfaces of both the bullet and rifle barrel. When the bullet is fired great heat will be generated and the oxygen layers of both barrel and nickel surfaces will be forced into intimate contact with the nickel. The conditions are thus favourable for the production of oxide. This is supported by the presence of oxide on the non-contact areas on which powder has been deposited. The powder at the temperature of the explosion would provide a more permanent and close source of heat than the passage of hot gases. The presence of the powder would thus assist the combination of the adsorbed oxygen layer and the nickel atoms. In the case of the lathe-polished specimens the abrasive pressure is small, and the presence of benzene probably reduces the temperatures considerably when large areas are in question. Although certain amounts of nickel oxide may be formed, the conditions are obviously not as favourable as in the other cases. This may explain the absence of nickel oxide in observable quantities in the lathe-polished and hand-polished layers.

	Block polish (A.)	Lathe polish (A.)	Contact surface (A.)	Non-contact surface (A.)
First halo	$1{\cdot}8 \pm 0{\cdot}04$	$1{\cdot}86 \pm 0{\cdot}04$	$1{\cdot}83 \pm 0{\cdot}04$	$1{\cdot}87 \pm 0{\cdot}03$
Second halo	$1{\cdot}0 \pm 0{\cdot}04$	$1{\cdot}08 \pm 0{\cdot}04$	$1{\cdot}11 \pm 0{\cdot}02$	$1{\cdot}12 \pm 0{\cdot}02$

The question of oxide in polish layers on metal has been considered by Dobinski[6]. In the above table it will be noticed that the spacings corresponding to the haloes obtained from the block-polish and lathe-polish layers are of the order of those obtained by Dobinski when he polished nickel under benzene. Dobinski obtained spacings of 1·99 A. and 1·16 A., and was of the opinion that these spacings were characteristic of a surface free from oxide. In the present experiments the contact and non-contact layers which give the spacings tabulated have been shown to consist largely of nickel oxide. There is a slight variation from Dobinski's spacings, and in view of this his opinion may be correct. On the other hand, no trace of nickel oxide was observed in the block-polish or lathe-polish layer when its composition was investigated, whereas its pattern has spacings very similar to those obtained from contact and non-contact layers. Indeed, the question of diffuse haloes on nickel seems a very open one. Plessing[7] has described polishing under various conditions, and his results show a large variation from Dobinski's. It would appear that nickel under conditions not yet determined can give haloes corresponding broadly to one or other of the Bragg-spacing values hitherto obtained. The results as summarized by Plessing are given below.

Plane spacings (A.)	Metal powder (A.)	Raether (A.)	Dobinski (A.)	Plessing (A.)
2·03 (111) 1·76 (200)	1·90	2·24	1·99	$2{\cdot}22 \pm 0{\cdot}07$
1·24 (220) 1·06 (311)	1·15	1·26	1·16	$1{\cdot}27 \pm 0{\cdot}04$

§4. ACKNOWLEDGEMENT

I wish to record my gratitude to Professor G. P. Thomson, F.R.S., for his unfailing help and encouragement throughout the work.

REFERENCES

(1) French, R. C. *Proc. Roy. Soc.* A, **140**, 637 (1933).
(2) Kirchner, F. *Ergebn. exakt. Naturw.* **11**, 112 (1932).
(3) Cochrane, W. *Proc. Roy. Soc.* A, **166**, 228 (1938).
(4) Hopkins, H. G. *Phil. Mag.* **21**, 820 (1936).
(5) Bowden, F. P. and Hughes, T. P. *Proc. Roy. Soc.* A, **160**, 575 (1937).
(6) Dobinski, S. *Phil. Mag.* **23**, 397 (1937).
(7) Plessing, E. *Z. Phys.* **113**, 36 (1939).

THE STRUCTURE AND ORIENTATION OF SILVER HALIDES

By H. WILMAN,
Applied Physical Chemistry Laboratories, Imperial College, London

Received 30 *October* 1939. *Read* 4 *April* 1940

ABSTRACT. An electron-diffraction study has been made of the reaction of halogen vapours with the {111}-twinned silver films prepared by condensation on hot rocksalt cleavage faces, and with similar films reduced to single-crystal structure by heat treatment.

Owing to the persistence of the silver twin spots and lines in the patterns from extensively attacked films, it seems likely that both these types of diffraction are to be interpreted as due to twinning alone, and not to presence of crystal surfaces parallel to octahedral planes. It also seems unnecessary to postulate extensive junction regions of the differently oriented lattices in the films. Accurate determinations of the lattice constants both of the twinned and of the single-crystal silver films have been made, and the values found are effectively identical with the x-ray value for normal pure silver in massive form.

Graphite was used as the most suitable comparison standard. It is pointed out that Finch and Fordham's gold comparison standard must have had the normal lattice constant, and that the large discrepancies in alkali-halide lattice constants found by them are connected with the nature of the specimens.

The silver halides formed had the normal structures, although the chloride crystals were often, and the cubic iodide crystals always, twinned on {111} planes. Strong spots due to secondary scattering were present in many of the patterns. The lattice constants of the unchanged silver and of the halide crystals were determined for several specimens and were found to be substantially constant and of normal value.

Rapid attack by concentrated halogen vapour always yielded a randomly disposed layer of halide crystals. Slower attack by halogens diluted with air yielded halide layers with strong 2-degree orientation relative to the original silver lattice, even when complete conversion of the film to halide resulted. Intermediate rates of attack produced a small proportion of crystals with 1-degree orientation as well as the crystals with 2-degree orientation.

§1. INTRODUCTION

In a previous study[1] it was shown that the electron-diffraction reflexion and transmission patterns yielded by thin silver films condensed in vacuo on to rocksalt cleavage faces kept at about 200° C. were in agreement with Menzer's[2] view of the {111}-twinned structure of such films. It was shown further that the detached films tended to develop local curvature about axes parallel to the film plane in the directions of the cubic axes and the cube-face diagonals of the part of the film which formed a lattice parallel to that of the rocksalt substrate, and that

such a film became entirely converted into a normal single-crystal lattice having the latter orientation on being heated in vacuo for several minutes at about 500° C.

An account is now given of an electron-diffraction study of the modes of growth of crystals of silver chloride, bromide and iodide, formed by the attack of silver films of both the twinned and single-crystal types of halogen vapours. One of the main objects of these experiments was to throw further light on the structure and surface form of the silver films before and after heating in vacuo.

§ 2. PREVIOUS DATA ON THE CRYSTAL STRUCTURE OF SILVER CHLORIDE, BROMIDE AND IODIDE

X-ray diffraction results[3]. Silver chloride and silver bromide have been shown[4], [5], [7], by means of X-ray powder photographs, to possess a cubic structure of the sodium-chloride type, in agreement with the holohedral cubic symmetry shown by crystals in the corresponding minerals cerargyrite (horn silver) and bromyrite. In connexion with the interpretation of the ionic conduction of electricity in silver bromide, Wagner and Beyer[14] measured the density of silver bromide at 20° and 410° C. and found from X-ray powder photographs that the lattice structure was unchanged in type, the ratio of the length of the cubic axis at 397° C. to that at 20° C. being $1{\cdot}021_4$.

Silver iodide is known to exist in three forms. A cubic structure, of zinc-blende type, was reported by Wilsey[4] (1921), and this was confirmed by Davey[5]. Aminoff[6] pointed out that the crystals in the mineral iodyrite were known to have hexagonal symmetry, and that this was confirmed by Laue patterns obtained from a cleavage sheet with the beam normal to (001), the cleavage plane, or to (100). Wilsey[4] (1923) studied the powder photographs from samples of silver iodide precipitated from the nitrate solution by potassium iodide, and concluded, from the variation of intensity of diffractions common to both proposed structures and from the presence of lines not common to both, that the cubic structure (zinc-blende type) and hexagonal structure (wurtzite type) occurred together in varying proportions; the only difference between the packing of the tetrahedral atom groups in the two structures being analogous to that between cubic and hexagonal close packing of spheres. This was confirmed by Barth and Lunde[7], who found that pure cubic samples could be obtained from silver iodide precipitated from solution and dried at 100° C.; and that samples of the powdered mineral were never purely hexagonal but always contained some of the cubic form, the difference between the powder patterns of the two forms being well marked. Kolkmeijer, van Dobbenburgh and Boekenoogen's[9] spectrometer measurements from (001) and (101) faces of natural iodyrite crystals yielded $d_{001}=7{\cdot}490$ A., $d_{101}=3{\cdot}508_6$ A., i.e. $a=4{\cdot}585_6$ A. and $c/a=1{\cdot}633_4$, from which the density was calculated to be 5·681 g./cm^3.

Bloch and Möller[10] found that the pure cubic form was obtained by powdering any coarse crystalline preparation, e.g. that solidified from the melt; and that the

pure hexagonal form occurred as the mineral and was obtained by dissolution of any silver-iodide preparation in concentrated KI solution followed by precipitation of the silver iodide by dilution, or by precipitation from HI or $Hg(NO_3)_2$ solution, or by allowing silver iodide to solidify from the melt with no further handling such as powdering or reheating. Samples precipitated from KI and $AgNO_3$ solution always yielded mixed cubic and hexagonal forms from which the diffractions 400, 331 and 620, characteristic only of the cubic form, were obtained besides those due only to the hexagonal form. The fraction with cubic form was increased by using strongly diluted solutions and large excess of $AgNO_3$, and Bloch and Möller concluded that in precipitation the cubic silver iodide is formed primarily, but at once partly dissolves in excess KI and later is reprecipitated, on washing, as the hexagonal form; so that the cubic form can never be obtained pure by precipitation. Both the cubic and hexagonal forms were converted into the same new form at about 146° c., and the x-ray patterns from different samples at various temperatures above 165° c. were the same, but the determination of the structure from them was not carried out. The remarkable fact was established that the proportions of the hexagonal and cubic forms in any sample were effectively unchanged by once heating to above the transition point, 146° c., and then cooling again (e.g. the pure cubic form remained cubic and the pure hexagonal form remained hexagonal), although the patterns yielded above the transition point were all similar. Initially pure hexagonal samples yielded a small fraction of the cold-cubic form on cooling after having been heated to near the melting point (which is at 555° c.) for several hours, but no hexagonal form was detected in pure cubic samples which had undergone the same treatment. By heating pure cubic and pure hexagonal samples for six weeks at $125° \pm 5°$ and $135° \pm 5°$ c. it was found that the hexagonal form is the stable one between 146° c. and about 135° c., and the cubic form is stable below 135° c. This was also confirmed by the action of iodine on silver mirrors deposited on thin glass rods, although their actual temperature during iodization was rather doubtful; reaction above 146° c. yielded the pure low-temperature cubic form of silver iodide, while just below the transition temperature both the cubic and hexagonal types were obtained, and the specimens formed below 135° c. were always pure cubic. Finally, Bloch and Möller also found that sublimed silver iodide consisted mostly of mixed cubic and hexagonal forms, though one sample was pure hexagonal.

Kolkmeijer and Hengel[11] found that silver iodide was precipitated from solution in the cubic form if Ag^+ ions were in excess, and in hexagonal form if I^- ions were in excess. The hexagonal form was prepared by dissolving silver iodide in KI solution, filtering, diluting, and washing the precipitate until no I^- ions could be detected by their reaction with $AgNO_3$. They calculated the parameter u of the hexagonal structure to be $0{\cdot}37_1$, i.e. near the ideal value of $\frac{3}{8}$.

Strock[12] obtained Hull-Debye-Scherrer photographs of silver iodide at a temperature near the melting point of tin, i.e. 230 to 250° c. Using a silver

pattern for calibration he determined the lattice constant of the cubic high-temperature silver iodide, and concluded from the intensities of the diffractions that with two molecules per unit cell the iodine atoms must form a body-centred cubic pattern ($I-I=2{\cdot}18$ A.) while the two silver atoms in the unit cell must be equally distributed among 42 possible places, not all of which, however, have the same coordination number.

Helmholz[13] used powder, Laue and oscillation methods, and concluded that at $-180°$ C. the structure of the hexagonal form is essentially that of wurtzite with the ideal parameter value for regular tetrahedral arrangement, while at room temperature the silver ions appeared to be distributed at random among the four positions tetrahedrally surrounding the ideal position.

The lattice constants of silver chloride, bromide and iodide calculated from the X-ray patterns are of special interest for comparison with the results described later; for this reason they are shown in table 1.

Electron-diffraction results. Natta[15] obtained electron-diffraction patterns from films of the silver halides precipitated by the alkali halide from silver hydroxide solution on thin collodion films and washed with distilled water. He found that freshly prepared specimens yielded normal ring patterns of the halide, but that a silver bromide film which had been exposed to light for a very long time yielded also diffractions characteristic of silver.

Trillat and Motz[16] have also studied the action of light and of electrons on silver halides by means of X-ray diffraction and electron diffraction. The pure silver-bromide specimens used for electron-diffraction examination were prepared in red light by complete reaction of thinned silver leaf floated on bromine water or KBr solution saturated with bromine, or by precipitation from an ammonia solution. Specimens of the first type yieded initially only a pattern of diffuse scattering, but after the specimens had remained $\frac{1}{4}$ to $\frac{1}{2}$ hour in the path of the beam a ring pattern due to crystalline silver bromide was visible, sometimes with indications of a weak preferred orientation with {111} planes parallel to the film surface. No silver diffractions were observed even after the electron beam had passed through the film for two hours, although a few seconds' exposure sufficed to solarize the plates used for recording the patterns. Specimens exposed to visible or ultraviolet light and then examined by electron diffraction also yielded the same phenomenon of crystallization, without any production of crystalline silver. That this crystallization of the bromide was not due to heating caused by the electron beam was shown by its not being affected by heating of the specimens between 150° and 300° C.; nor was it due to slow spontaneous crystallization in the film. It was shown that although no pattern due to crystalline silver was obtained from the films after exposure to light or electrons, the exposed parts became grey and insoluble in sodium hyposulphite after the film had been developed by being floated on a bath of developer, and that these parts then yielded an electron-diffraction pattern due to crystalline silver. Silver-bromide specimens prepared either in darkness or in the light, by evaporation of the solution in

Table 1. Lattice constants of AgCl, AgBr and AgI obtained by x-ray diffraction

Reference	Diffraction method *	Type of material	Lattice constants (A.) AgCl (cubic)	AgBr (cubic)	AgI Blende type (cubic)	AgI Wurzite type (hexl.)[1] a	AgI Wurzite type (hexl.)[1] c	AgI Wurzite type (hexl.)[1] c/a	AgI High temp. form (cubic)
4a (1921)	P	Precipitate from $AgNO_3$ by Khalides	2·78	2·89	6·53				
		Fused salt	—	,,	—				
5	P	? (not stated)	2·76	2·89	6·53				
4b (1923)	P	Precipitate from $AgNO_3$	5·540	5·768	6·493	4·593	—	1·633	
		Fused salt cooled slowly or quenched in cold water			,,	,,		±·008	
6	L	(001) cleavage plates from artificially prepared crystals				4·59	7·53	(1·64)	(cubic)
	P	,, ,,				4·57	7·50	(1·64)	
7	P	Precipitated and dried at 100° C.	5·545	5·755	6·491				
8	P	Precipitated from $AgNO_3$ washed, dried		2·89					
9	P	? (" prepared ")			—	—	—	—	
	S	(001) and (101) faces of natural crystals				$4·585_6$	7·490	$1·633_4$	
10	P	Any coarse-grained AgI powdered			6·49				
		Precipitated from $AgNO_3$ or AgI solution in KI, HI or $Hg(NO_3)_2$				4·59	7·53	(1·64)	
11	P	Precipitated from KI solution				$4·58_0$	$7·49_4$	$1·63_7$	
12	P	Powdered hexagonal crystals (?) [10]			$6·47_3$				5·034 ±·006
13	P	" As described by Bloch and Möller "			—	—	—	—	
	L	Thin plate-like crystals grown from Ag solution in HI				—	—	—	
	O	(001) cleavage face, as cleaved or after grinding				4·59	7·52	1·64	
						—	—	—	

* P, powder ; L, Laue ; S, spectrometer ; O, oscillation.

ammonia on a cellulosic substrate film, immediately yielded well defined patterns due to the bromide crystals, sometimes showing {111} orientation on the substrate. The patterns did not change with longer exposure to the electron beam, and no silver pattern was observed. Similar results were obtained with silver-iodide films prepared by the action of iodine vapour on silver leaf; thus both low-temperature forms were obtained together, and no crystalline silver was detectable until the film had been developed in a developing bath.

In later experiments Trillat and Merigoux[17] used films (<1000 A. thick) of silver bromide and gelatine, some parts of which yielded, initially and also on prolonged exposure to the beam, Hull-Debye-Scherrer patterns of the bromide, from which the lattice constant was estimated to be 5·78 A. In this case, however, when those parts of the film which did not yield initially a silver-bromide pattern were exposed to the electron beam for from three to five seconds, diffraction rings due to crystalline silver appeared, and these were attributed to solarization due to the action of the electron beam, rather than to changes corresponding to the formation of a latent image. Long exposure (1½ hours) of these parts to the electron beam caused no further change in the pattern. Films of the emulsion which had been exposed and developed yielded, besides the gelatine haloes, only silver patterns which did not change on prolonged exposure to the beam. It was concluded that the parts of the emulsion which are effective in producing photographic action on exposure to light or electrons contain imperfectly formed crystals of silver bromide.

§3. EXPERIMENTAL PROCEDURE

Silver films a few hundred angstroms thick were examined by electron diffraction when freshly prepared in the manner described previously[1] and mounted on a nickel gauze support. Such films and those subsequently heated in vacuo to produce a homogeneous single-crystal lattice were again examined after exposure to halogen vapour in air for various times and at different concentrations of the vapour. The Finch[18] type of electron-diffraction camera was used with a camera length and accelerating potential of about 50 cm. and 60 kv. respectively. The silver films were allowed to react with the halogen vapour in the presence of air in a large glass flask (of about 2 litres capacity) at room temperature and in daylight. Bromine and iodine were available in a sufficiently pure elementary state, slight warming of the iodine being necessary to produce conveniently a suitable concentration of vapour. Chlorine was prepared by oxidation of hydrochloric acid with potassium permanganate, and was partially dried before use by being passed over calcium chloride.

§4. THE GROWTH AND STRUCTURE OF SILVER CHLORIDE ON SILVER FILMS

The silver films, after removal from the rocksalt substrates on which they were deposited, and again after subsequent heating in vacuo, yielded electron-diffraction patterns similar to those published previously,* showing that the silver

* Reference (1), figures 3, 8, 10, 11 and figure 14 respectively.

lattices were of fairly perfect orientation and were of {111}-twinned single-crystal type and of normal single-crystal form respectively.

In the first experiments silver films of each type were placed for two minutes in a chlorine-air mixture with rather high percentage of chlorine, and it was found that under such conditions the attack was so rapid that a white silver-chloride layer was then noticeable. The films yielded electron-diffraction patterns which showed that the silver chloride crystals were relatively large and mainly disposed at random with respect to each other and the silver lattice, even though the latter was particularly homogeneous in orientation. Thus the unchanged silver yielded a pattern of strong spots limited to the broad Laue zones due to the cubic axis which was normal to the film plane and nearly parallel to the beam; while the silver chloride yielded a strong and sharp ring pattern, and in some cases the presence of very short intensified arcs or spots on these rings (similar to those of figure 1) indicated that some of the chloride crystals had grown in a definite orientation relative to the silver lattice, namely, with the (001) plane parallel to

Table 2. Main types of diffraction patterns obtained from silver films after attack by chlorine-air mixtures

Pattern type	Approximate percentage of Cl_2 in air	Total period of attack (min.)	Diffraction pattern from films after attack	Orientation of the AgCl crystals on the silver lattice
1	80	2	Same type as figure 8; Ag pattern weaker, unaltered in type; strong sharp rings due to AgCl	Mainly random
2	5	½	As figures 1 and 2; Ag pattern slightly weaker; very short arcs, almost spots, due to AgCl, and sometimes with faint AgCl rings; considerable secondary scattering of the strong Ag diffracted beams by the underlying layer of AgCl	Mainly very strong :— (i) entirely (figure 1) or mostly (figure 2) : {001} AgCl parallel to (001) Ag ⟨100⟩ AgCl parallel to [110] or [1$\bar{1}$0] Ag; (ii) sometimes also(figure 2): {111} AgCl parallel to (001) Ag ⟨1$\bar{1}$0⟩ AgCl parallel to [100] or [010] Ag
3	5	1½	As figures 1 and 2, but Ag pattern weaker and AgCl pattern stronger	,, ,,
4	5	7	As figure 3; no Ag pattern left; strong AgCl pattern	,, ,,
5	5	1 to 5	Cf. figure 4; Ag spots more or less strong; AgCl patterns similar to types 2, 3 and 4 but containing additional spots.	Strong orientation as (i) (sometimes with (ii) also) above, the additional spots showing also presence of AgCl crystal lattices in the {111} twin positions to this

the (001) silver plane (i.e. the film-plane) and [100] parallel to [110] or [1$\bar{1}$0] of the silver lattice.

Further experiments were made with a slower rate of attack, i.e. by the use of a chlorine-air mixture containing roughly 5 per cent of chlorine, and with periods of attack of from 2 seconds to 10 minutes successively for the same film. Fifteen silver films were examined, each at various stages of attack, and the different types of diffraction patterns yielded by them after typical periods of reaction are

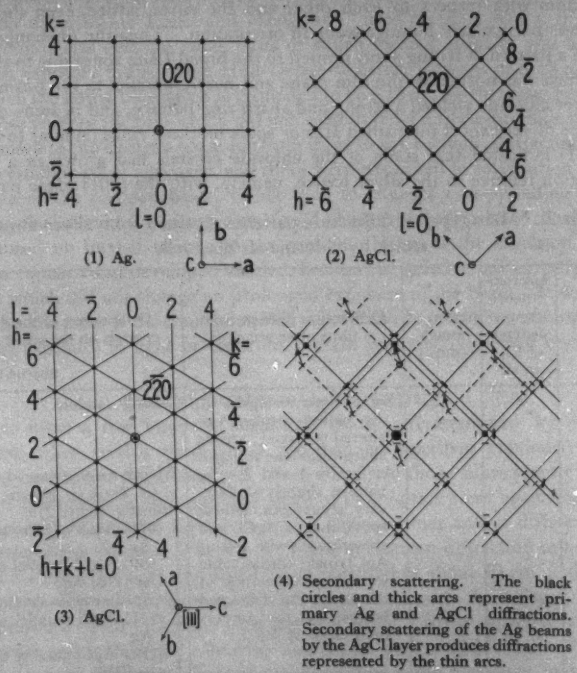


Figure 5. Analysis of the Ag–AgCl patterns.

described in table 2. In all cases the spot pattern due to the unchanged silver, when present, was closely similar to that obtained from the initial film, apart from progressive weakening as the conversion to chloride proceeded; and in particular the twin spots and the lines joining them to the normal diffractions, if present initially, remained unaltered in form and relative intensity.

Interpretation of the patterns. Figures 1, 2 and 3 were obtained with the beam normal to the film-plane, or nearly so. The main component patterns which are recognizable in them are shown diagrammatically in figure 5, in which

patterns (1), (2) and (3) are drawn to the correct scale relative to each other, while (4) is enlarged to show the detail of the secondary-scattering effects.

Figure 5 (1) represents the spot pattern due to the initial silver crystal in the above setting, the twin spots being omitted. The lines in figure 5 (1) indicate the even-order Laue zones due to the two axial lattice rows which are normal to the beam, the distance apart of the zones being $2\lambda L/a$, where λ is the electron wavelength, L is the camera-length, and a the lattice constant of the silver. At the intersections of these zones occur the corresponding cross-grating spot diffractions. Since the crystals are thin, the Laue zones due to the atom rows parallel to the beam are very broad, and only the zero-order zone comprising the region near the central spot need be considered in connexion with figures 1 to 3. Thus, if we take as the c axis the cube axis parallel to the beam, and if the a and b axes are disposed as shown below figure 5 (1), the spot pattern in the zero-order c zone ($l=0$) is of the form shown in figure 5 (1), since h and k must be even when l is even.

Besides this pattern in figure 1, there occurs the spot pattern represented by figure 5 (2); the positions and spacings of the spots show that the silver-chloride crystals had grown in a single orientation with respect to the silver lattice, with a cube face parallel to that of the silver lattice in the film plane, and the other cube faces at 45° to cube faces of the silver lattice. With the axes a, b, c disposed as shown below figure 5 (2), the 220 diffraction of the chloride lies very close to the 020 spot due to the silver, on the same radial line from the central spot, since the length of the cubic axes of the chloride is very nearly $\sqrt{2}$ times that of the silver lattice; and similarly all the other diffractions in the region near the central spot lie either close to silver diffractions or near the centre of the squares formed by the latter.

In figure 5 (4) these two patterns of spots are shown superposed (drawn to a larger scale, with the difference in size exaggerated slightly for greater clarity), the silver diffractions and those of the chloride being represented by black spots and short thick arcs respectively. The thinner arcs show the positions of the more conspicuous of the diffractions resulting from secondary scattering of the strong silver diffracted beams by the underlying chloride layer. Thus, for example, if the 020 diffracted beam due to the silver is again partially diffracted as it passes through the underlying chloride layer, it will give rise to a pattern of the chloride diffractions based on its intersection with the screen as central spot, instead of on the true central spot corresponding to the undeflected part of the primary electron beam; and the less deviated, i.e. stronger, diffractions associated in this way with the 020 silver diffraction are distinguished in the diagram by means of the slightly displaced network of dotted lines and arrows. The whole set of diffractions of this type due to all the silver diffracted beams occurs quite strongly in the patterns in figures 1 and 2. Actually a chloride layer was formed on each side of the silver films, and the electron beam passed first of all through one such layer, whose diffracted beams must also, in part, undergo further scattering by the underlying silver and the second chloride layer. It can be shown in the same manner as that

outlined above that the first part of this process will yield secondary diffractions, some of which are in the same positions as those shown in figure 5 (4), while some are in positions intermediate between or near two neighbouring diffractions of that group. Such diffractions are not clearly visible in the patterns and are apparently too weak to be observed owing to their being further weakened by diffraction in the lower chloride layer. The second of the above processes would not give rise to any new diffracted-beam positions if the chloride crystals were all oriented exactly parallel to each other and yielded a pattern of spots, but owing to the small deviations from this ideal case, which cause slight arcing in the chloride patterns, the resulting pattern of secondary scattering will consist for the most part of small blurred regions surrounding the main chloride arcs, such as are to be seen in figures 1 and 2.

In figure 2 there occur, besides the silver-chloride pattern shown in figure 5 (2), that of figure 5 (3) and a similar pattern rotated by 30° with respect to this about the central spot. This pattern indicates that some of the silver-chloride crystals were oriented with a {111} plane parallel to the film-plane cube-face of the silver lattice, and with a cube-face diagonal parallel to one or other of the two silver cube-axes in the film-plane. We are here only concerned with the zero-order Laue zone of diffraction from the [111] axis parallel to the beam; hence it is convenient to construct the chloride pattern expected for one of these settings by drawing the Laue zones due to the cube axes, because the latter are all steeply inclined to the beam at the angle $\cos^{-1}(1/\sqrt{3})$, i.e. 54° 44′, so that over the region concerned their Laue zones may be taken as practically straight lines which are normal to the projections of these axes on the screen and have a spacing $\lambda L/a \sin 54° 44'$, where a is the cube axis spacing. There are four orientations of the chloride crystals on the silver lattice which are geometrically equivalent to that represented by the a, b, c axial projections drawn below figure 5 (3), the three others being obtained from this by successive rotations of the diagram by 90° about the [111] axis, which is normal to the paper; thus presumably, on the average, chloride crystals are formed in all four orientations with the same degree of frequency. When the chloride crystals occurred in the same specimen in these and the cube-face orientations, as in figure 2, a characteristic subsidiary pattern of secondary diffractions was formed in a manner similar to that outlined above, and can be seen in figure 2; it is especially noticeable in the region near the central spot and innermost normal diffractions. When the stage of complete conversion of the film to chloride was reached, the silver spot pattern and the pattern due to secondary scattering were absent and only the strong chloride pattern remained, as is shown in figure 3.

Figure 6 shows diagrammatically the above orientations of the chloride crystals with respect to the silver lattice. The circles represent the atomic positions in a (001) silver plane parallel to the mean film plane, and the black spots and crosses represent the silver and chlorine ions respectively, in a parallel plane of the superposed silver-chloride lattice, one atom being arbitrarily shown as situated

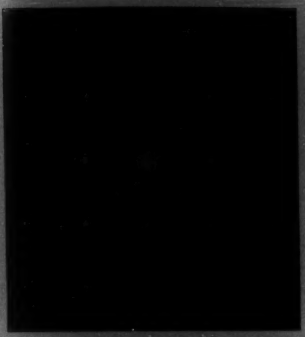


Figure 1. Ag+AgCl.

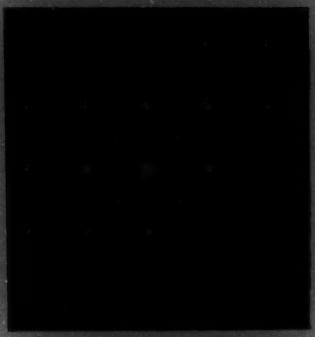


Figure 2. Ag+AgCl.

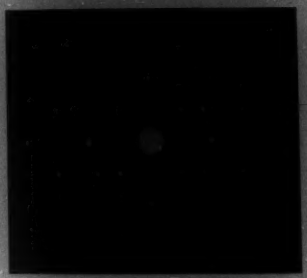


Figure 3. AgCl.

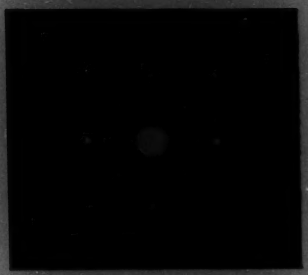


Figure 4. Ag+AgCl.

Figure 7. Ag+graphite.

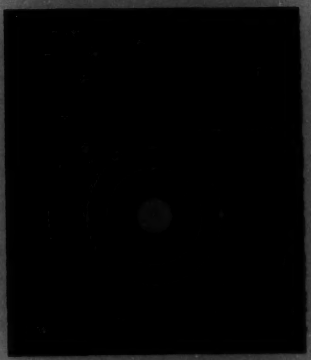


Figure 8. Ag+AgBr.

To face page 332

directly above an atom in the silver (001) plane in order to show the closely similar spacings of parallel atom rows in these two planes. In the case of the {111} planes of the chloride the alternate planes contain all silver atoms or all chlorine atoms in a plane hexagonal arrangement. The upper surface of the silver films after deposition on the rocksalt substrates is probably in some cases a cube face, and after subsequent removal of the substrate and heating in vacuo both surfaces appear to be cube faces; thus the atom planes represented by figure 6 may possibly be the actual junction planes of the two lattices in these cases.

The chloride pattern of type 5 (table 2) contains diffraction spots in addition to the spot pattern of figure 5 (2), forming with the latter a pattern similar to that yielded by the initial unheated silver films. The unsymmetrical intensity-distribution of the spots in figure 4 shows that the corresponding mean film-plane must have been not quite normal to the beam. All the additional diffraction

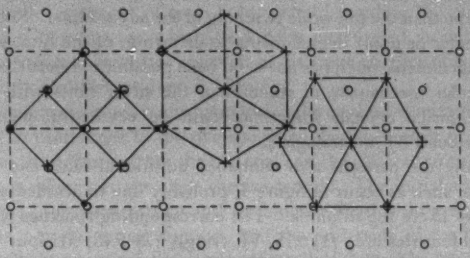


Figure 6. The observed AgCl orientations on Ag, shown by atomic planes parallel to the mean film-plane.

spots in figure 4 lie on the normal Hull-Debye-Scherrer ring positions of the chloride, e.g. 111, 200, 311, 222; and at first sight it seemed possible to explain the occurrence of these spots as due to curvature of the normal oriented chloride layer corresponding to figure 5 (2) about its two cube-face-diagonal directions in the film-plane. The amount of curvature away from the main average setting required to cause the appearance of diffractions of 111 type lying on the 111 ring is roughly $90°-\xi$, where ξ is the angle of inclination of the film plane away from the setting normal to the beam, i.e. ξ is here of the order of 10°. Since very little curvature of the unchanged part of the silver film is indicated from the pattern, it would be necessary to suppose that the chloride film had become largely detached and curled away from the silver film. This would not, moreover, explain the appearance of the spots, or rather short arcs, on the 200 and 111 rings at their intersections with the radii which pass through the spots whose indices are $\pm 2, \pm 2, 0$ in the main chloride pattern, nor can they be explained by assumption of any other kind of curvature of normal chloride crystals.

On the other hand, these diffractions, as well as those already described, can readily be explained as due to {111} twinning of chloride crystals whose orientations are represented by figure 6 (i), assuming also a relatively small amount of curvature, probably due mainly to curvature of the silver film, about directions parallel to the mean film-plane and to the cube-face diagonals of the initial lattice on which the twins are developed, exactly in the way postulated for the interpretation of the similar patterns from the initial unheated silver films[1]. Since the patterns do not contain diffraction lines joining up the diffractions as in the patterns from the initial silver films, it is possible that the twinning may be apparent instead of real, i.e. the polycrystalline chloride layer may contain normal untwinned crystals in the same orientations as the lattices formed by the above set of twins and their common initial lattice. If the twinning is real, however, as seems likely, it must be concluded, from the absence of the diffraction lines and other satellite diffractions, that the chloride crystals had grown in the characteristic orientation relative to the silver lattice, but had developed a {111}-twinned structure of a coarser type than the fine-scale twinning of the silver films. Similar patterns, arced to a greater or lesser degree owing to distortion of the film, were obtained from several specimens, but it has not so far been possible to relate their occurrence to any particular mechanism of reaction of the silver films with the halogen. That silver-chloride crystals sometimes twin on octahedral planes has been observed previously by mineralogists.

Occasionally the patterns also contained additional faint and more diffuse rings and arcs, such as occur in figure 4, probably due to accidental traces of an impurity, most likely the sulphide. The corresponding spacings (in angstroms) and relative intensities are: (1) 6·15, VF (ring); (2) 5·48, M (four arcs at 90°); (3) 4·36, F (four arcs at 45° to those of (2) and on silver 200 radii); (4) 2·75, M (four arcs on same radii as (2)); (5) 2·42, F; (6) 1·726, F (four arcs on same radius as (3)).

Determinations of lattice constants. The accurate determination of the lattice constants of both the initial silver films and the chloride and silver lattices after partial reaction of the films with chlorine was of considerable importance not only for comparison with the x-ray values obtained from larger crystals and in connexion with the mutual orientation of the lattices, but also to determine whether a change in the silver lattice constant could be detected after reaction of the films with chlorine.

The most reliable and convenient means of determining the lattice constants is that suggested by Finch and Fordham[19], i.e. by comparison with the length of the basal axis in graphite, since this depends only on the carbon-carbon spacing in the (001) planes and therefore can be assumed to be constant, independent of either crystal-thickness or possible presence of absorbed gas or other slight impurity. Finch and Fordham obtained double-shutter photographs of gold-leaf and graphite (Aquadag) ring patterns, and calculated from the measurements of these that, assuming the gold lattice constant to be the same as that found by means of x rays, i.e. 4·070 A., the length of the graphite basal axis a was 2·458 A.

Shortly afterwards Finch and Wilman[20] obtained the value 2·726 for the axial ratio (c/a) of graphite. These measurements were considerably more accurate than those of the previous X-ray determinations of the lattice constants of graphite, but there remained some doubt of the constancy of the lattice constant of gold in thin films, especially in view of the discrepancies found between previous X-ray values for lattice constants of several of the alkali halides and Finch and Fordham's values determined by electron diffraction with gold leaf as a standard. Since then, however, two sets of more accurate X-ray determinations have been carried out, and these confirm the electron-diffraction results. Thus Hofmann and Wilm[21] have obtained $a=2{\cdot}455\pm0{\cdot}002$ A. (taking a NaCl equal to 5·626 A. although 5·628 A seems to be the more generally accepted value), $c=6{\cdot}69\pm0{\cdot}01$ A. and $c/a=2{\cdot}72_5$; and Trzebiatowski[22] has found $a=2{\cdot}456_3$ A., $c=6{\cdot}695_6$ A., and $c/a=2{\cdot}726_2$ (taking $a_{Ag}=4{\cdot}07758$ A.). Since the electron and X-ray values for a of graphite can reasonably be assumed to be identical, we may now say, conversely, that Finch and Fordham's results show that their gold foil had effectively the normal lattice constant of pure gold, notwithstanding the results obtained by some other investigators which have been interpreted as indicating a considerable real variation in lattice constant in different samples of gold leaf, which were apparently of different degrees of purity.

In the present case it was convenient to use the method of mixtures; hence this was adopted in preference to the two-specimen double-shutter method, in order to ensure that the patterns from the material and the calibration substance should be recorded at effectively the same camera-length. The initial silver films, after removal of the rocksalt substrate and washing by floating on a water bath, were picked up on a nickel gauze support, and the underside of the latter was then temporarily allowed to touch the surface of a colloidal graphite solution (Aquadag) so that when the film had been dried in vacuo a thin layer of graphite was obtained in contact with the silver film. The sharp and strong 110 graphite ring in the composite electron-diffraction transmission pattern was alone used for reference, since the 100 ring was less intense and too small for accurate measurement, while the remaining $hk0$ rings were too faint. In the silver patterns which contained short 311 arcs owing to distortion of the film, these arcs always lay nearly exactly on the graphite 110 ring, as may be seen in figure 7. Taking the 110 graphite spacing, $d_{c.110}$, to be $1{\cdot}228_1$ A., it follows directly that the silver lattice constant a_{Ag} is close to 4·073 A. Measurements of the diameters of the graphite 110 ring and of the silver $hk0$ diffraction spots yielded a more accurate value. The plates were measured by a microscope made by the Cambridge Instrument Company (drum graduations corresponded to 0·001 cm.; readings by estimation to 0·0001 cm.), and two readings were taken for each diffraction, traversed in opposite directions, the settings being usually reproducible to 0·001 cm. or less. The radii R obtained from the four readings taken at opposite ends of each diameter were converted to values of R' given by $R'=R-R\delta$, where δ is a tabulated variable depending on R/L (L being the camera-length), such that the corresponding net-

plane spacings d are given accurately by the equation $d=\lambda L/R'$. In calculating the silver lattice constant a_{Ag} we have, therefore,

$$a_{Ag}=R_{c.110}\times 1{\cdot}228_1\,(h^2+k^2+l^2)^{\frac{1}{2}}/R'_{hkl} \text{ angstroms.}$$

When the 200 and 220 silver-diffraction radii were used, three plates yielded by one unheated (i.e. twinned) silver specimen gave the values $a_{Ag}=4{\cdot}079$, $4{\cdot}075$ and $4{\cdot}084$ A.; and two plates from another specimen gave $a_{Ag}=4{\cdot}080$ and $4{\cdot}080$ A. The mean of these values, $4{\cdot}079\pm 0{\cdot}003$ A., is practically that found with X rays for silver, namely $4{\cdot}077$ A. No change in lattice constant was detected after similar silver films had been heated in vacuo and thereby converted into normal single-crystal lattices.

The composite silver-graphite films, after exposure to chlorine-air mixtures, yielded transmission patterns in which the clear and sharp diffractions due to silver, silver chloride and graphite were readily recognizable. Measurement of the diametral distances of about six diffractions in each of the patterns, and use of the diameter of the graphite 110 ring as a standard for comparison, yielded the first series of values in table 3, while patterns from various specimens which were

Table 3. Calculated values of a_{Ag} and a_{AgCl} from various photographs (values relating to different specimens are given on different lines)

Reference basis (A.)	AgCl pattern type (cf. table 2)	a_{AgCl} (A.)	a_{Ag} (A.)
$d_{c.110}=1{\cdot}228_1$	2	5·549	4·080
,,	2	5·548	$4{\cdot}077_5$
$a_{Ag}=4{\cdot}077$	2	5·544	—
	2 and 3	5·551, 5·548 *	—
	2	5·545	
	5	5·546	—
	Mean values	5·547	4·079

* The same specimen after further reaction with chlorine.

not mixed with graphite gave the lower series of values of a_{AgCl} in table 3. Thus it appears that the lattice constant of the silver films is unaltered by partial conversion of the silver to chloride and re-examination in vacuo in the diffraction camera; and the silver-chloride lattice-dimension found from the diffraction patterns obtained after various degrees of reaction of the films is also effectively constant and independent of whether or not the chloride develops a {111}-twinned structure. The mean value of a_{AgCl} is 5·547 A., in good agreement with the previous X-ray results given in table 1.

§ 5. THE GROWTH AND STRUCTURE OF SILVER BROMIDE ON SILVER FILMS

The patterns obtained from the silver films after reaction with bromine vapour (in presence of air) were in general similar to those obtained from the silver, silver-chloride specimens. Thus, rapid attack with bromine-air mixtures containing mostly bromine yielded a yellowish layer of random crystals of silver

bromide (cf. patterns of type 1, table 2), as shown by the ring pattern in figure 8, which was obtained with the beam nearly normal to the film-plane. The spots in figure 8 are due to the unchanged part of the (initially heat-treated) silver film, and the perfection of orientation of the silver lattice in this region of the film is shown by the fact that the spots lie on a well-marked circular zero-order Laue zone which passes through the central spot and is associated with the lattice rows along the cube-axis direction normal to the film-plane.*

Dilute bromine-air mixtures reacted with the silver films to yield mainly highly oriented bromide crystals with the same orientations relative to the silver lattice as were found in the case of the chloride (cf. table 2, pattern-types 2 and 3). When only the orientation (i) was present, many of the *hk*0 bromide spots fell almost exactly on the *hk*0 silver diffractions, and the extent to which the reaction had proceeded had to be judged by comparing the intensities of diffractions of this set with those of the remaining diffractions. Figure 9, for example, shows a pattern which was due to a film containing a small proportion of unchanged silver, as is shown also by the fact that the faint twin spots can still be seen which were associated with the spots of 200 type in the pattern from the initial silver film. Owing to the fact that the bromide lattice-dimension is almost exactly equal to $\sqrt{2}$ times that of the silver, secondary scattering merely tends to even out the spot-intensities in patterns of the above type. When bromide crystals occurred in all the equivalent orientations of types (i) and (ii), the additional diffractions caused by secondary scattering were prominent, however, as may be seen in figure 10, which was obtained from a silver film that had been heat-treated in vacuo before being subjected to partial reaction with bromine. Silver films which were not heat-treated initially, i.e. had a {111}-twinned structure, yielded, after reaction with bromine, patterns such as figure 11, in which the pattern due to the unchanged silver with its twin spots and related streaks can be seen to be exactly similar to the patterns from the initial silver films of this type. No evidence was found for the occurrence of any {111}-twinning of the bromide crystals, such as was often met with in the chloride patterns (cf. figure 4). The only observed case of abnormal diffractions was in a pattern from a film completely converted into bromide, a faint ring (spacing 2·464 A.) occurring nearly midway between the 200 and 220 rings of the pattern, and bearing twelve superposed spots of moderate intensity which lie on the same radii as the twelve spots on the 220 bromide ring; and several faint rings are visible nearer to the central spot in the same photograph (corresponding to spacings of 6·02, 3·49, 3·31, 3·00, 2·79, 1·86 and 1·75 A.). These diffractions probably belong to the same set as was observed in a few silver, silver-chloride patterns, and were probably due to a trace of silver sulphide.

The lattice constants of the silver bromide calculated from the patterns yielded by three specimens are shown in table 4, where the first series was obtained by comparison with graphite in the way outlined in connexion with the chloride;

* Only the central part of the 3¼ × 4¼-in. plate is shown in figure 8, however.

and the lower value relates to the specimen (free from graphite) yielding the pattern figure 8, assuming the silver lattice constant to be 4·077 A., i.e. the X-ray value. In the latter case the crystal-breadth and pattern-resolution were sufficiently great to show that the silver diffractions lay very slightly nearer the central spot than the associated bromide rings; hence the bromide lattice constant was slightly less than $\sqrt{2}$ times the silver cube-axis length, i.e. $\sqrt{2}\times 4{\cdot}077$, or 5·766 A.

Table 4. Calculated values of a_{Ag} and a_{AgBr} from several Ag–AgBr specimens

Reference basis (A.)	AgBr pattern type (cf. table 2)	a_{AgBr} (A.)	a_{Ag} (A.)
$d_{c.110}=1{\cdot}228_1$	2	5·766	4·080
	2	5·763	4·080
$a_{Ag}=4{\cdot}077$	1 (figure 8)	5·763	
	Mean values	5·764	4·080

Measurements of the silver 400—400 spot separation and the bromide 440 ring-diameter in figure 10, and also of the corresponding diameters at 90° to this direction, yielded the values for the radii: $R_{Ag\ 400}=2{\cdot}458_5$ and $2{\cdot}458_2$ cm., and $R_{AgBr\ 440}=2{\cdot}464_5$ and $2{\cdot}462_5$ cm. respectively, whence the means are $2{\cdot}458_3$ and $2{\cdot}463_5$ cm. and if $a_{Ag}=4{\cdot}077$ A. it follows that $a_{AgBr}=5{\cdot}763$ A. This value is nearer Wilsey's[4] value than that of Barth and Lunde (cf. table 1), and is probably more accurate than either.

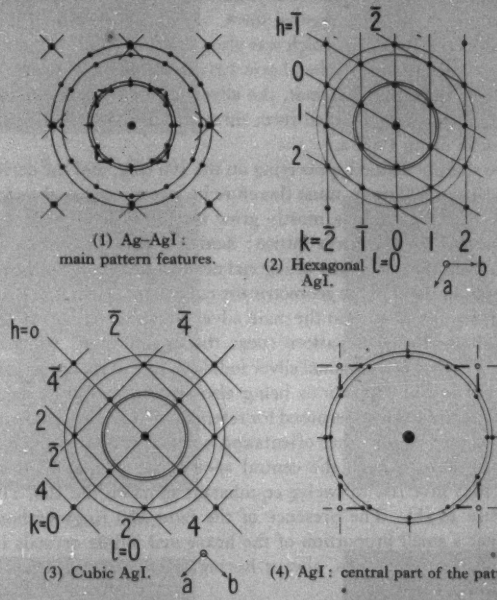
§ 6. THE GROWTH AND STRUCTURE OF SILVER IODIDE ON SILVER FILMS

As in the case of the chloride and bromide, it was found that a rapid attack of the silver films by iodine vapour (in presence of air) caused growth of an iodide layer of randomly disposed crystals. Figure 12 shows the Hull-Debye-Scherrer type of pattern from such a specimen which had been entirely converted into silver iodide, the sharpness of the rings indicating a mean particle-diameter of at least 200 A.

Slower attack of silver films by weaker mixtures of iodine vapour with air yielded iodide layers which possessed a well-marked orientation relative to the silver lattice. A film which had undergone partial attack by being placed for thirty seconds over a flask containing a few flakes of iodine slightly warmed yielded the pattern figure 13 when nearly normal to the beam. The pattern due to the unchanged silver in figure 13 is still much stronger than that of the iodide layer, and shows the twin spots and streaks which are apparently exactly similar to those yielded by the film before reaction with the iodine. After further attack the specimens yielded patterns similar to figure 13, except that the silver pattern became weaker and the silver-iodide diffractions became progressively stronger until only the iodide pattern was present, i.e. the film was completely converted into iodide.

The very thin iodide layers produced by an exceedingly slight attack yielded patterns closely similar in type to figure 13, but in which the iodide pattern was very faint relative to the silver pattern.

Interpretation of the patterns. The interpretation of the patterns required special care owing to the kind of polymorphism of the silver iodide, as well as the simultaneous occurrence of different crystal-orientations in the iodide layer and the possibility of twinning and curvature of the crystals. It was found possible, however, to reach clear and definite conclusions as to the film structure by the



(1) Ag–AgI: main pattern features.

(2) Hexagonal AgI.

(3) Cubic AgI.

(4) AgI: central part of the pattern.

Figure 14. Analysis of the Ag–AgI patterns.

following process of analysis applied to the patterns such as figure 13 from the oriented specimens which, though more complicated than the ring patterns (figure 12), contain correspondingly much more detailed information about the film structure.

Figure 13 was obtained from a partially attacked unheated silver film with the film-plane roughly normal to the electron beam; and figure 14 (1) shows diagrammatically the relative positions of the diffractions near the central spot, for the purpose of discussing the pattern. The strong silver *hk*0 spots are represented in figure 14 (1) by the large black spots through which pass the diagonal lines

which are ascribed to {111} twinning of the silver crystals combined with the appreciable curvature of parts of the silver film about the cube-edge and cube-face diagonal directions in the mean film-plane(1). Considering the remaining part of the pattern it is observed that the innermost ring has twelve spots superposed on it equidistantly round its circumference, and there are arcs on the same radii of the pattern, on the two prominent rings which pass near the silver 200 diffractions. Measurement shows that these three rings are clearly the 100, 110 and 200 diffractions of the hexagonal form of silver iodide. The first and third of these at least cannot be due also to crystals of the cubic form; hence much of the iodide had grown in the hexagonal form. From the measured diameters of these rings and of the 400 ring, which was also clear and could be measured with good accuracy, the length a of the basal axis was calculated by comparison with the radial distance of the silver 200 spot, the silver lattice constant being assumed to be 4·077 A. The values obtained from three specimens were 4·586, 4·580 and 4·583 A., the mean being 4·583 A.

The twelve sharply defined spots lying on the 100 ring, and the corresponding spots on the 110 and 200 rings, must therefore be interpreted as showing that the crystals of the hexagonal iodide mostly grow in two closely-followed preferred orientations relative to the silver lattice; namely, with the vertical axis [001] normal to the silver film-plane, the hexagonal networks of atoms in the basal plane (001) being in one or other of the geometrically equivalent orientations represented by figures 6 (ii) a and b relative to the main silver substrate lattice. Figure 14 (2), for example, shows the spot pattern (near the central spot) which would be obtained from a crystal of hexagonal silver iodide in one of these two orientations, the a, b and b' Laue-zone positions being also shown (straight lines), and the rings present in figure 14 (1) superposed for reference. The pattern corresponding to a hexagonal crystal in the other orientation is obtained by rotating figure 14 (2) by 30° about an axis through the central spot perpendicular to the diagram; thus together they give rise to twelve equidistant spots on the 100, 110 and 200 rings as in figure 14 (1). The presence of the faint $hk0$ rings without any hkl rings shows that a small proportion of the hexagonal iodide crystals had grown in other azimuthal orientations, always having their (001) axes normal to the film-plane however.

Just outside the innermost ring (hexagonal 100) is another faint ring with strong arcs on it as shown in figure 14 (1). This ring has a radius close to that required for the hexagonal 002 diffraction but is also in the position required for the 111 diffraction from crystals of the low-temperature cubic form of silver iodide. Thus it may be seen that the positions of the arcs lying on it, and of the peculiar streaks passing through them, are very reminiscent of those on the 111 ring of the initial silver patterns; cf. figure 14 (1). A cubic-iodide crystal oriented with a cube-face, (001) for example, parallel to the silver film-plane, would yield the spot pattern shown in figure 14 (3). The straight lines in figure 14 (3) represent the a and b Laue-zone positions, and the circles shown in figure 14 (1) also are

Figure 9. Ag+AgBr.

Figure 10. Ag+AgBr.

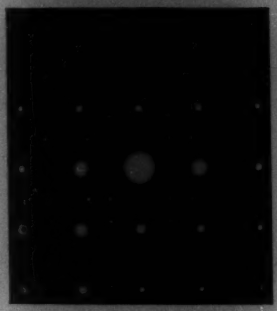


Figure 11. Ag+AgBr.

Figure 12. AgI.

Figure 13. Ag+AgI.

Figure 15. Ag+AgI+NaCl.

To face page 340

marked for reference. If a zinc-blende type of structure is assumed for the cubic iodide, the $hk0$ diffractions with $h+k+l$ equal to $4n+2$, where $n = 0, 1, 2, 3$ etc. have a very small structure factor, while those with $h+k+l$ equal to $4n$ have a very large one; hence diffractions such as 200, 420 etc. will be very weak and those like 220, 400 etc. very strong. The apparent absence of the ± 200 and ± 020 spots in figure 13 is therefore to be expected, and the strong spots with indices $\pm 2, \pm 2, 0$, though coinciding with four of the twelve spots of hexagonal-110 type, evidently account for the observed strengthening of these four arcs relative to the others. No spots are observed in figure 13 in the positions corresponding to the expected ± 400 and ± 040 diffractions of figure 14 (3), but this is explained by the rapid radial decrease in intensity of the iodide pattern owing to the decrease in atomic scattering factor and to the relative weakness of the iodide pattern. The expected diffractions of 400 type are present with moderate intensity in patterns such as figure 15 obtained from films which had undergone a more extensive attack by iodine and therefore contained a larger proportion of iodide; and faint spots of 200 type also occur there. Thus the above interpretation is satisfactorily confirmed. Comparison of the radial distances of the cubic iodide 400 spots with those of the silver 200 and 220 spots in figure 15 yielded the value 6·486 A. for the lattice constant of the cubic silver iodide.

The arcs on the second ring (the cubic 111 ring) and the streaks passing through them must therefore be regarded as closely analogous to those of the silver patterns, and as arising from the {111}-twinning of the cubic iodide crystals combined with the curvature of parts of the iodide layer about the same axes as those about which the original silver film was curved; this curvature is thus probably due almost entirely to the initial local curvature of the silver films.

This group of diffractions is indicated more clearly by figure 14 (4), which is drawn to twice the scale of figures 14 (1), (2) and (3). There appear to be no clear satellite spots of the other group found in the silver patterns, although continuous diffraction streaks are visible in figure 15, which is evidently from a film in which the local curvatures are more extensive than in that yielding figure 13; hence it seems that the twinning is probably not on such a fine scale as in the initial silver films. Faint diffuse streaks also appear to extend outwards a short way from the twelve 100 spots yielded by the hexagonal iodide, being due similarly to the local curvatures of the film.

Since the position of the hexagonal 110 ring is the same as that of the cubic 220 ring, the twelve spots on it may arise partly from cubic crystals in the (111) orientations on the film, as the corresponding spots for the chloride and bromide are found to do. Occurrence of other spots in the pattern in related positions show that some cubic-iodide crystals do probably occur in these two equivalent (111) orientations. The atomic arrangement in these planes is the same as that in the hexagonal (001) planes, and these atomic planes take up the same two geometrically equivalent dispositions in relation to the silver substrate lattice.

Having established clearly that both the cubic and hexagonal forms of silver

iodide are produced under the conditions used, and that these do not yield prominent satellite diffractions, we can now use the ring pattern of figure 12 to obtain a reliable value for the axial ratio of the hexagonal iodide. The *hk*0 diffraction rings could be easily identified in figure 12, whence the remaining rings also were readily indexed. Table 5 gives the mean value of the radii R,

Table 5. Analysis of the AgI ring pattern of figure 12.

Intensity	Ring radii R (cm.)	R' (cm.)	d (A.)	Hexagonal indices *hkl*	Cubic indices	Hexagonal form		Cubic form; a (A.)
						a (A.)	c/a	
MS	$0{\cdot}564_6$	$0{\cdot}564_5$	3·959	100		4·581		
S	$0{\cdot}598_3$	$0{\cdot}598_1$	3·736	002	111	(*c*=7·472)	(1·630)	6·471
M	$0{\cdot}636_6$	$0{\cdot}636_5$	3·511	101			1·642	
F	$0{\cdot}685_7$	$0{\cdot}685_6$	3·259	—	200			6·518
F	$0{\cdot}819_4$	$0{\cdot}819_3$	2·727	102			1·637	
S	$0{\cdot}976_3$	$0{\cdot}976_2$	2·290	110	220	4·580		6·474
MF	$1{\cdot}055_7$	$1{\cdot}055_5$	2·117	103			1·639	
MF	$1{\cdot}127_8$	$1{\cdot}127_6$	1·983	200		4·590		
MS	$1{\cdot}145_3$	$1{\cdot}145_0$	1·952	112	311		(1·625)	6·473
VF	$1{\cdot}165_5$	$1{\cdot}165_3$	1·917	201			—	
M	1·194	$1{\cdot}193_7$	1·872	004	222	(*c*=7·488)	(1·634)	6·485
VF	$1{\cdot}267_4$	$1{\cdot}267_1$	1·764	202			—	
F	1·317	$1{\cdot}316_8$	1·697	104			1·643	
MF	1·378	$1{\cdot}377_5$	1·623	—	400			6·492
MF	$1{\cdot}438_2$	$1{\cdot}437_7$	1·554	203			1·635	
M	$1{\cdot}496_3$	$1{\cdot}495_8$	1·494	120	331	4·575		6·513
F	1·512	$1{\cdot}511_4$	1·479	121			—	
F	$1{\cdot}601_7$	$1{\cdot}601_0$	1·396	122 105				
M	$1{\cdot}690_4$	$1{\cdot}689_5$	1·323	300	422	4·592		6·480
F	$1{\cdot}736_5$	$1{\cdot}735_5$	1·287	123			1·643	
M	$1{\cdot}791_9$	$1{\cdot}790_9$	1·247	006 302	333 511	(*c*=7·482)	(1·633)	6·483
VVF	$1{\cdot}906_7$	$1{\cdot}905_4$	1·173	124			(1·641)	
MF	$1{\cdot}944_8$	$1{\cdot}943_2$	1·150	220	440	(4·600)		6·504
M	2·039	2·037	1·097	130	? 531	(4·578)		

The values in brackets were not used in taking the means in view of the inaccuracy due to their coincidence with strong cubic diffractions or to the form of their indices being unfavourable for the determination of axial ratios.

Mean a_{cubic}=6·489 A. (from d_{400} we have a=6·492 A.).

Mean $a_{\text{hexl.}}$=4·583 A. (assumed from other patterns where the silver was used as reference).

Mean c/a=1·640±0·003; thus c=7·521 A.

the reduced radii R', and the plane-spacings d, calculated from the relation $d=\lambda L/R'$, where λL is determined by assuming the *hk*0 spacings to be those obtained from the value of a found above, i.e. 4·583 A. From these spacings and the indices *hkl* given in the next column the values of the axial ratio were calculated, the mean of the most reliable (from the 101, 102, 103, 104, 203 and 123 diffractions) being 1·640±0·003. The 200 and 400 diffractions of the cubic iodide occur independently, while the strong 111, 220, 311, 331 etc. diffractions practically

coincide with diffractions from the hexagonal crystals. The radius of the 200 ring is too small for much accuracy in measurement, but the measurement of the 400 ring is more reliable and yielded $a_{\text{cubic}}=6{\cdot}492$ A., in good agreement with the mean value 6·489 A. obtained from all of the cubic diffractions, although these values varied rather widely, presumably owing to the effect of the superposed hexagonal diffractions.

§ 7. CONCLUSION

In a previous paper[1] it was found that the electron-diffraction patterns from the unheated silver films appear to be wholly explicable in terms of the {111}-twinned structure of the deposits and the local curvatures present in the thin films, and do not require the postulation of an octahedral form of surface. The present experiments have shown that the only change in the silver patterns, when the twinned films are attacked by halogen vapours, is a general decrease in intensity relative to the halide pattern. Thus it seems improbable that the irrational diffraction spots and lines in the silver patterns are due, even in part, to the presence of octahedral faces, whether at the film-surface or as crystal boundaries within the film. The fact that the orientation assumed by the halide crystals on such films is the same as that on the heat-treated (single-crystal) films is also of interest in this connexion, although the preferred relative orientation of two lattices is not necessarily different for different substrate boundary planes. The shape of the lower surface of the films is presumably that of the substrate.

It also seems unnecessary to suppose that junction regions of the differently oriented lattices form any appreciable part of the films, though the diffraction lines must be attributed to thin layers of the twinned lattices. The twin spots in the patterns correspond to a certain group of the points which divide into sixths the diagonals of the identity-cells of the combined reciprocal lattice. All the remaining, much weaker, spots of this group can be attributed to secondary scattering of strong diffracted rays in the film. The strongest beams correspond to the *hkl* diffractions of a silver lattice parallel to the rocksalt substrate lattice, and even if these only are scattered appreciably by the other twin lattices they will suffice to give the whole series of the diffractions in question, with about the intensities observed. As has already been mentioned, strong secondary-scattering effects were often found in the patterns from composite silver, silver-halide films, and similar secondary diffractions in composite films have indeed been described previously[22].

The recent precision determinations[21, 22] of the lattice constants of graphite are of importance in the use of colloidal graphite as the most reliable and convenient comparison standard for the determination of lattice constants by electron diffraction. They also provide confirmation of the value of the axial ratio calculated by Finch and Wilman[20], and proof that the gold used as a standard by Finch and Fordham[19] had indeed the normal lattice constant, as they assumed. The large discrepancies between the X-ray lattice constants and those found by Finch and Fordham[19] for some of the alkali halides are undoubtedly real and appear

to be connected with the fact that the halide specimens were prepared by condensation from the vapour. Thus the silver films such as those used in the present experiments, if not thoroughly washed after the rocksalt substrate has been dissolved away, bear a thin layer of salt recrystallized from the solution, and yield spot patterns (analogous to figure 1) containing both silver and sodium-chloride diffractions; and from these, when the normal lattice constant as found above is taken for silver, the NaCl lattice constant is calculated to be 5·630 A., i.e. the normal value (5·628 A.) to within the possible error. Figure 15 shows such spots.

The very close agreement between the lattice constant found for the silver films and the X-ray value for massive silver was to be expected, at least in the case of the heat-treated (i.e. single-crystal) films, since the thickness of the films was probably of the order of 500 A. It is of greater interest, however, that the lattice constant of the unchanged silver in films partly attacked by halogens was found to have the normal value also; and those found for the halides were also constant to within the possible error. These facts, combined with the complete absence of any disarrangement of the partially attacked silver lattice, may be interpreted as showing that the reaction occurs entirely at or close to the surface of the silver, beneath the layer of halide already formed.

The course of the reaction may be visualized somewhat as follows. Halogen atoms incident on the silver surface will tend to react with silver atoms there, the reaction being exothermic. The reaction evidently consists of ionization of the silver and halogen atoms, which then aggregate together to form silver halide lattices, i.e. small crystals one or two atoms thick and of limited lateral extent. If the ions in this thin initial layer can take up positions in contact with each other and also rest in the hollows between the neighbouring atoms in the substrate surface, they will do so, to occupy the positions of least potential energy. The small, nearly two-dimensional, crystals of the thin halide layer may therefore be expected to grow in definite preferred orientations if there is a close similarity of spacing (with a difference less than about 15 per cent) of the atoms in a densely populated row of the halide lattice and those in one of the lattice rows of the substrate surface, since a short atom-row of this type in the halide lattice might be built up first and act as a nucleus for the formation of a crystal. In the present experiments the first detectable pattern due to a halide layer has, indeed, shown practically complete two-degree orientation of the component crystals, except in cases of rapid attack of the film; and the halide crystals must have been at least 100 A. in extent in directions parallel to the film-plane. The succeeding stages of the attack are less clear, and perhaps consist in the chlorine atoms diffusing between the neighbouring halide crystals and under them along the interface between the halide and the silver and reacting there with the silver, the resulting ions being then added to the underside of the halide crystals without disturbing the orientations of the latter. It may be mentioned that for all three halogens the radii of the ions are much greater than those of the atoms, and also the closest silver-halogen distance is slightly greater for ions than for atoms; hence it would be geometrically

more favourable for penetration of the halogen as atoms than as ions, followed by ionization of a halogen atom simultaneously with that of a silver atom. It might also be possible for halogen ions to be formed on the upper surface of the halide layer and then to migrate through the halide crystal to the under side by a process of crystal-movement such as that suggested by Bernal[24], but this seems to be incompatible with the perfection of the halide lattice indicated by the patterns even from films completely converted to halide (figure 3), and with the good agreement in the lattice constants calculated for different films.

Strong two-degree orientation of the crystal lattices of the reaction product on a single-crystal substrate has also been observed in electron-diffraction studies of the oxidation of copper[25, 26], zinc-blende[27, 28], stibnite[29], palladium[30] and molybdenite[31] single crystals; and in several x-ray studies of a similar type, such as the oxidation of iron crystals.

The lattice constants found for the silver halides are probably somewhat more accurate than the best previous x-ray determinations. They are as follows: $a_{AgCl}=5{\cdot}547\pm0{\cdot}002$ A., $a_{AgBr}=5{\cdot}764\pm0{\cdot}002$ A., $a_{AgI\ cubic}=6{\cdot}489\pm0{\cdot}005$ A., and $a_{AgI\ hexl.}=4{\cdot}583\pm0{\cdot}002$ A., $c/a=1{\cdot}640\pm0{\cdot}003$. The axial ratio for the hexagonal iodide agrees well with many of the x-ray values and with that found crystallographically by von Zepharovich[32], i.e. 1·63920.

With regard to the orientations assumed by the silver-halide crystals relative to the unchanged silver, the general observation that rapid attack results in randomly disposed crystals whereas slow attack yields strongly oriented layers is readily interpreted as being due to the difference in rate of liberation of energy in the reaction. Thus, for all three halides the reaction is strongly exothermic, and this means that the more rapidly it takes place the more intense is the atomic disturbance at the film surface.

Another point of interest is that the thin chloride crystals take up the same orientations whether they develop a {111}-twinned structure or not. This appears to indicate that in the case of silver films condensed on hot rocksalt cleavage faces it may not be necessary for the initial silver deposit to have a twinned structure for the observed orientation relative to the rocksalt to be produced.

The cubic halide crystals all take up the same orientation relative to the silver substrate lattice in spite of their widely different lattice constants. In the (001) orientation the *a* and *b* halide cube axes are at 45° to those of the silver in the film plane, and the AgBr lattice has then practically the same identity periods as the parallel ones in the silver lattice, but this is only approximately so for the chloride and even less so for the iodide. The halide spacing parallel to a cube-axis direction in the film plane is $a/\sqrt{2}$, and for AgCl, AgBr, AgI this is 3·922, 4·076 and 4·588 A. respectively; thus the differences from the silver spacing (4·077 A.) are $3\frac{3}{4}$, 0·002 and $12\frac{3}{4}$ per cent respectively. The order of decrease in similarity of spacing is the same as that of the observed tendency of the crystals to develop a twinned structure, which is also least for AgBr and most for AgI. Probably the conditions inducing twinning are similar to those operating in the case of the silver deposits

during their formation, although, as was stated above, the absence of twinning in many of the chloride specimens seems to show that twinning need not necessarily occur in the region immediately adjacent to the substrate. In the case of the initial silver layers, however, Kirchner and Cramer [33] have published a reflexion pattern from a very thin deposit, showing that at least in this case the initial layers had a marked twin structure.

In the (111) orientation the halide crystals take up a position with a cube-face diagonal parallel to one of the cube axes of the silver substrate, and the differences between their identity spacings and those of the silver in these directions are thus the same as in the above case of the (001) orientation ; whereas in other directions no simple relationship between the atomic positions of the adjacent lattices is apparent. This suggests, as has already been stated, that a preferred two-degree orientation of crystals grown on a single-crystal substrate depends fundamentally on there being a similarity of spacing (with a difference less than about 15 per cent) in *one* row of closely spaced atoms in each lattice. Thus, in the initial deposition of atoms on the substrate, such an atom row (not necessarily a long one) of the deposit lattice would naturally tend to be built up most readily and parallel to the corresponding row of the substrate, since a single atom of the deposit would tend to occupy a potential trough or hollow between projecting parts of atoms of the substrate surface ; and other deposited atoms, tending to do likewise, could do so and yet occupy a normal position of packing side by side with each other. Further additions of atoms in normal deposit-lattice positions at the side of such a row would then most likely take place so that a small crystal would be formed, with one of its planes of closest atomic packing oriented parallel to the mean substrate surface ; or if a densely packed plane occurred in which atom-rows in another direction could simultaneously fit in with the periodicity of the substrate lattice, then the plane containing this and the former atom-row would probably tend to grow parallel to the mean substrate surface. Besides considerations of this sort, however, the relative concentrations of anions and cations, and the form and regularity or otherwise of the surface, must have some part in determining which orientations are taken up by the deposit lattice. Thus, if the anion concentration is greatest, the initial plane of the deposit crystal may be one in which the anions predominate or occur alone, as in the orientation (ii) of figure 6 ; and if there are many cations at the surface during formation of the initial halide layer, a plane containing a close-packed arrangement of the two together such as occurs in the (001) orientation, represented in figure 6 (i), may then be built up most readily on the substrate surface.

§ 8. ACKNOWLEDGEMENTS

It is a pleasure to thank Prof. G. I. Finch for many helpful discussions during the work, which was carried out in his laboratories. Acknowledgements are due to Monsieur M. Macquestiau, who carried out part of the experimental work in the Francqui Laboratory at Brussels University during Prof. Finch's tenure of the Francqui chair (1937–8). The Conseil d'Administration of the University and

that of the Fondation Francqui are also to be thanked in this connexion for the facilities made available in Brussels. Finally, the author wishes to thank the Council of the Royal Society for the award of a Moseley Research Studentship, during the tenure of which the work has been carried out.

REFERENCES

(1) Goche, O. and Wilman, H. *Proc. Phys. Soc.* **51**, 625 (1939).
(2) Menzer, G. *Naturwissenschaften*, **26**, 385 (1938); *Z. Kristallogr.* **99**, 378 & 410 (1938).
(3) The earlier x-ray literature is conveniently listed by R. W. G. Wyckoff in 'The structure of crystals', New York (1931); Supplement (1935).
(4) Wilsey, R. B. *Phil. Mag.* **42**, 262 (1921); **46**, 487 (1923).
(5) Davey, W. P. *Phys. Rev.* **19**, 248 (1922).
(6) Aminoff, G. *Geol. Fören. i. Stockholm Förhandl.* **44**, 444 (1922). (Abstract in *Ber. Phys. dtsch. Ges.* **3**, 827 (1922)); *Z. Kristallogr.* **57**, 180 (1922).
(7) Barth, T. and Lunde, G. *Norsk Geol. Tiddskr.* **8**, 281 (1925). (Abstract in *Chem. Zbl.* **1**, 2771 (1926)); *Z. phys. Chem.* **122**, 293 (1926).
(8) Koch, P. P. and Vogler, H. *Ann. Phys., Lpz.*, **77**, 495 (1925).
(9) Kolkmeijer, N. H., van Dobbenburgh, W. J. D. and Boekenoogen, H. A. *Proc. K. Akad. Sci. Amst.* **31**, 1014 (1928).
(10) Bloch, R. and Möller, H. *Z. phys. Chem.* A, **152**, 245 (1931).
(11) Kolkmeijer, N. H. and van Hengel, J. W. A. *Z. Kristallogr.* **88**, 317 (1934).
(12) Strock, L. W. *Z. phys. Chem.* B, **25**, 441 (1934); B, **31**, 132 (1935).
(13) Helmholz, L. *J. Chem. Phys.* **3**, 740 (1935).
(14) Wagner, C. and Beyer, J. *Z. phys. Chem.* B, **32**, 113 (1936).
(15) Natta, G. *Trabayos del IX Congreso Internacional de Quimica Pura y Aplicada, Madrid*, **2**, 177 (1934).
(16) Trillat, J. J. and Motz, H. *J. Phys. Radium*, **7**, 89 (1936).
(17) Trillat, J. J. and Merigoux, R. *J. Phys. Radium*, **7**, 497 (1936).
(18) Finch, G. I. and Wilman, H. *Ergebn. exakt. Naturw.* **16**, 353 (1937).
(19) Finch, G. I. and Fordham, S. *Proc. Phys. Soc.* **48**, 85 (1936).
(20) Finch, G. I. and Wilman, H. *Proc. Roy. Soc.* A, **155**, 345 (1936).
(21) Hofmann, U. and Wilm, D. *Z. Elektrochem.* **42**, 504 (1936).
(22) Trzebiatowski, W. *Roczn. Chem.* **17**, 73 (1937).
(23) Finch, G. I. and Sun, C. H. *Trans. Faraday Soc.* **32**, 852 (1936).
(24) Bernal, J. D. *Trans. Faraday Soc.* **34**, 834 (1938).
(25) Thomson, G. P. *Proc. Roy. Soc.* A, **133**, 1 (1931).
(26) Yamaguti, T. *Proc. Phys.-Math. Soc. Japan*, **20**, 230 (1938).
(27) Yamaguti, T. *Proc. Phys.-Math. Soc. Japan*, **17**, 443 (1935).
(28) Aminoff, G. and Broome, B. *Nature, Lond.*, **137**, 995 (1936).
(29) Miyake, S. *Sci. Pap. Inst. Phys. Chem. Res., Tokyo*, **34**, 565 (1938).
(30) Fordham, S. and Khalsa, R. G. *J. Chem. Soc.* p. 406, 1939.
(31) Uyeda, R. *Proc. Phys.-Math. Soc. Japan*, **20**, 656 (1939).
(32) Von Zepharovich. *Z. Kristallogr.* **4**, 119 (1880).
(33) Kirchner, F. and Cramer, H. *Ann. Phys., Lpz.*, **33**, 138 (1938).

THE BOILING POINT OF SELENIUM

By M. de SELINCOURT, B.A.,
Physics Department, The National Physical Laboratory, Teddington

Received 6 *December* 1939

ABSTRACT. In the long gap between the boiling point of sulphur (444°·60 c.) and the freezing point of gold (1063° c.) there is no fixed point which can conveniently be used for the routine calibration of thermocouples in terms of the International Temperature Scale. The boiling point of selenium (about 690° c.) is suggested as a possible fixed point in this region, and as a first step towards enabling it to be used for this purpose the precise determination of its value has been undertaken.

The determination was carried out in a silica boiling tube similar to that normally used at the sulphur point, having a re-entrant tube fitted with a radiation shield. By connecting the apparatus to a reservoir the variation of the boiling point with pressure over the range 700 to 800 mm. of mercury was also determined. The measurements were made by means of three standard platinum *vs.* platinum-rhodium thermocouples calibrated at the freezing points of gold, silver and antimony. Samples of selenium from two sources, and two types of radiation shield, were employed and gave the same result.

The value found for the boiling point in terms of the International Temperature Scale was 684·8$\pm$0°·1 c. at normal atmospheric pressure, with a variation of 1°·08 c. per cm. of mercury.

§ 1. INTRODUCTION

In the region of temperature from the boiling point of sulphur (444°·60 c.) to the freezing point of gold (1063°·0 c.) inclusive, four fixed points are recognized by the International Temperature Scale. These comprise, as basic points, the two just mentioned and the freezing point of silver (960°·5 c.) and, as a subsidiary point, the freezing point of antimony (630°·5 c.).

It is often desirable to calibrate thermocouples by direct reference to fixed points, but of those named only the sulphur and gold points can be realized with the ease essential for routine work. For the former any suitable boiling tube may be used and for the latter the so-called "wire method", in which the "hot" ends of the thermocouple wires are joined by a short length of gold wire. If the e.m.f. is observed when the hot junction is subjected to a slowly rising temperature, a short halt, followed by the breaking of the wire, indicates the melting point.

The freezing points of antimony and silver, on the other hand, are considerably more troublesome to determine, since the necessity of protecting the metals from the atmosphere and of using substantial ingots involves elaborate apparatus.

It is clearly desirable to bridge the long gap between the sulphur and gold points for the purpose of routine calibrations. The freezing point of sodium

chloride (about 801° C.) is sometimes used, but it is hardly less troublesome to realize than the silver and antimony points, and its value is not established with the same certainty. A possible alternative is the boiling point of selenium, which is stated to be in the neighbourhood of 690° C. As a first step towards making this point available a precise determination of its value has been undertaken.

§ 2. PREVIOUS DETERMINATIONS

The only determinations of the boiling point of selenium since 1900 which have been traced are those of D. Berthelot[1] (1902) and of Preuner and Brockmoller[2] (1913).

Berthelot, using a method in which the temperature of a column of air immersed in the vapour of boiling selenium is deduced from its refractive index, obtained four values between 685° C. and 694° C., giving a mean of 690° C. He found a variation with pressure of 1° C. per cm. of mercury. Another experiment, using a platinum *vs.* platinum 10 per cent iridium couple calibrated at the zinc point (419°) and the gold point (1064°) and with its scale calculated from a logarithmic relation between temperature and e.m.f., gave the value 690°·6 C. at normal pressure. Preuner and Brockmoller, using a resistance thermometer, found that the boiling point at normal pressure lay between 686° C. and 689° C., and that the variation with pressure was about 1°·2 C. per cm. of mercury. These determinations do not permit a numerical value to be assigned with sufficient accuracy for the purposes of the International Temperature Scale.

§ 3. EXPERIMENTAL DETAILS OF PRESENT EXPERIMENTS

Two samples of selenium obtained from different sources were employed; the first, with which the bulk of the experiments were made, was supplied by Messrs. Hopkin and Williams of London, and was stated to be of a purity of 99·9 per cent, the impurities not being specified. The second was supplied by Messrs. Schering-Kahlbaum of Berlin, and was stated to be 99·95 per cent pure, the remainder being almost exclusively iron oxide. This selenium was stated to be free from sulphur and other impurities.

The experimental procedure was as follows:—

The sample was boiled in a silica tube 60 cm. long × 5 cm. bore, insulated with sil-o-cel powder packed in an aluminium furnace case, as shown in figure 1. To avoid the risk of super-heating the vapour it is necessary to concentrate the heating energy as near the bottom of the tube as possible; for this reason the heating element was constructed by winding brightray tape on five sides of a 2-inch cubical former made of silica rod. The basket thus formed was placed in a cavity hollowed out of a diatomite brick 12 by 9 by 4½ inches in size, and slipped over the base of the boiling tube, which was left projecting from the bottom of the furnace case for this purpose. Should a failure in the winding occur with this arrangement, it would be a simple matter to remove and replace the heating element without allowing the selenium to freeze, with the consequent risk of fracturing the boiling tube.

The temperature measurements were made with three platinum *vs.* platinum 10 per cent rhodium thermocouples, calibrated at the freezing points of antimony, silver and gold. The thermocouple was placed in a re-entrant tube of 6 mm. bore silica, supported from a water-cooled brass cap, which was waxed to the top of the boiling tube; this cap was also penetrated by a $\frac{1}{2}$-inch bore glass tube which could be connected either with the atmosphere, or with a 30-gallon reservoir maintained at an artificial pressure, the difference from atmospheric pressure being measured with a water manometer.

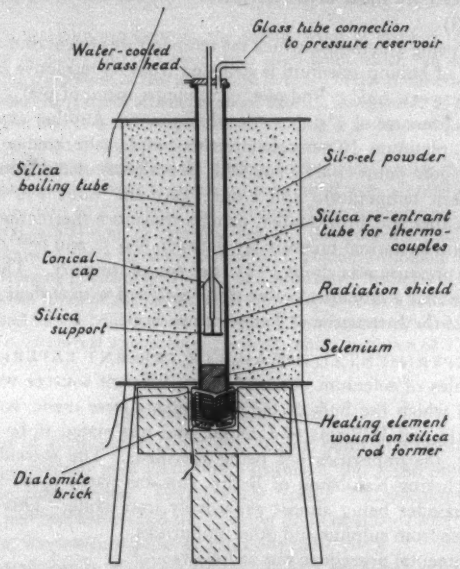


Figure 1. Selenium boiling-point apparatus.

The re-entrant tube terminated 8 cm. above the surface of the liquid, and was provided with a radiation shield and cap resting on silica supports. Two types of radiation shield were used; one consisted of a pythagoras tube of 2·5 cm. bore by 12 cm. long by 2 mm. wall with a conical cap of asbestos paper supported 5 mm. above the flat top of the radiation shield, and overlapping it by 5 mm. all round; the other of an alundum tube 3·0 cm. bore by 8 cm. long by 1 mm. wall with a cap, also conical, of the same material, resting directly on the castellated top of the radiation shield, and also overlapping by 5 mm. In both cases the radiation shield projected 2·5 cm. below the bottom of the re-entrant tube.

It is appropriate here to refer to the possible risk of poisoning arising out of the use of selenium. According to Mellor[3] poisonous properties are possessed

only by compounds of selenium, for example selenites and selenates, and not by the element itself. Nevertheless, since it is inevitable that the oxide should be formed in the process of boiling the element, it was thought advisable in the present investigation to conduct the experiments in a well-ventilated fume chamber.

§ 4. RESULTS

(1) A series of observations was taken with the apparatus connected to the reservoir at various pressures between 700 and 800 mm. of mercury, to determine the variation of the boiling point with pressure. The results of this determination, which are set out in table 1, indicate a variation with pressure of 1°·08 c. per cm. of mercury.

Table 1. Variation of boiling point with pressure

Pressure (mm. of mercury)	Boiling point (° c.)	Departure from straight line (° c.) sloping 1°·08 c. per cm.
715·4	680·1	+0·1
735·9	682·2	0·0
737·9	682·4	−0·05
765·7	$685{\cdot}4_5$	−0·03
789·0	688·0	+0·03
795·7	$688{\cdot}7_5$	+0·05
798·9	689·1	+0·06
804·8	$689{\cdot}6_5$	−0·03

(2) Two series of observations were taken with the apparatus open to the air, at atmospheric pressure between 745 and 770 mm., firstly with different thermocouples, and secondly with different radiation shields.

In these experiments it was found that, once boiling was established, the e.m.f. remained constant, apart from occasional bumping, to $\pm 0{\cdot}5$ microvolt. In addition, during the course of the observations, the energy supplied to the boiling tube was varied from 600 to 750 watts, which caused the condensation level to rise by some 10 cm., without affecting the e.m.f. by more than $\pm 0{\cdot}5$ microvolt, indicating that the vapour was not appreciably superheated. The position of the thermojunction could also be raised by 6 cm. without affecting the e.m.f. by more than this amount.

The results obtained with the different thermocouples and different radiation shields are set out below.

Table 2. Boiling point determined with different thermocouples

Couple	Pressure (mm. of mercury)	Temperature (° c.)	Boiling point at normal pressure (° c.)
S I	768·8	$685{\cdot}8_5$	$684{\cdot}9_0$
S II	765·7	$685{\cdot}4_5$	$684{\cdot}8_3$
P 27	767·6	$685{\cdot}7_3$	$684{\cdot}9_1$

Table 3. Boiling point determined with different radiation shields

Radiation shield	Pressure (mm. of mercury)	Temperature (° c.)	Boiling point at normal pressure (° c.)
Pythagoras, 12 cm.×2·5 cm. bore. Asbestos cap.	740·2	$682{\cdot}6_7$	$684{\cdot}8_1$
	749·9	$683{\cdot}7_4$	$684{\cdot}8_3$
Alundum, 8 cm.×3·0 cm. bore. Alundum cap.	766·2	$685{\cdot}5_5$	$684{\cdot}8_8$
	765·0	$685{\cdot}3_8$	$684{\cdot}7_8$
	773·0	$686{\cdot}3_5$	$684{\cdot}9_5$

It will be seen that the differences between the individual couples and between the mean values for the two types of radiation shield each fall within 0°·1 c., which is less than the estimated experimental error. The mean value for the boiling point at normal pressure is taken as 684°·8 ±0°·1 c.

(3) As a check on the purity of the selenium, a final series of observations was carried out with two couples on the second sample of selenium referred to above. The following results yield a mean indistinguishable from the value already given.

Table 4. Boiling point determined with different thermocouples

Couple	Pressure (mm. of mercury)	Temperature (° c.)	Boiling point at normal pressure (° c.)
S II	768·5	$685{\cdot}6_9$	$684{\cdot}7_8$
P 27	768·5	$685{\cdot}8_5$	$684{\cdot}8_9$
Mean	—	—	$684{\cdot}8_4$

§ 5. SUMMARY

The boiling point of selenium is suggested as a possible subsidiary fixed point on the International Temperature Scale in the gap between the boiling point of sulphur and the freezing point of gold, and as a first step its precise determination is undertaken.

The determination is carried out in a conventional boiling tube by means of three platinum *vs.* platinum rhodium thermocouples calibrated at the freezing points of antimony, silver and gold. The value is determined at various pressures between 700 and 800 mm. of mercury with samples of selenium from two different sources and with two types of radiation shield.

The value found for normal atmospheric pressure is 684°·8±0°·1 c., with a variation with pressure of 1°·08 c. per cm. of mercury.

REFERENCES

(1) Berthelot, D. *Ann. Chim.* (*Phys.*) **26**, 105, 142 (1902); *C.R. Acad. Sci., Paris*, **1**, 705 (1902).

(2) Preuner and Brockmoller. *Z. phys. Chem.* **81**, 144 (1913).

(3) Mellor, J. W. *A Comprehensive Treatise on Inorganic and Theoretical Chemistry*, **10**, 752. (London: Longmans, Green & Co., 1930.)

AN ALL-ELECTRIC CLOCK

By P. VIGOUREUX, D.Sc. and H. E. STOAKES

Received 28 *December* 1939. *Read* 9 *February* 1940

ABSTRACT. The synchronous motor usually employed for obtaining seconds' intervals from a 1 kc./sec. source controlled by a high-frequency crystal oscillator is replaced by three multi-vibrators whose controlled periods are 10, 100, and 1000 times that of the source respectively. On account of the extremely sharp rise of current in the multi-vibrator circuits, the intervals provided by the seconds multi-vibrator are equal to within 10 micro-seconds.

Each multi-vibrator consists of a single triode-pentode valve, injection from one stage to the next being effected on the suppressor grid of the pentode section, a method of injection which has proved superior to anode-circuit or grid-circuit injection. A check that the multi-vibrators are in step is obtained by using their outputs to charge a condenser through the coils of a cathode-ray oscillograph whose time base is synchronized by the 1 kc./sec. input.

Crystal-controlled clocks have in recent years proved in many ways superior to pendulum clocks[1], and are being installed in national laboratories and observatories. The frequency of the controlling quartz oscillator is in general of the order of 100 kc./sec.; the output from the oscillator controls electrical generators (reaction oscillators or multi-vibrators) of lower frequency, and, usually in two steps, the period is increased a hundredfold, viz., to 1 millisecond.

Hitherto it has been the practice at this stage to effect the further thousandfold increase in period by mechanical means: the 1 kc./sec. output drives a synchronous motor at, say, 10 r.p.sec., and through reduction gear the motor drives a wheel at 1 r.p.sec. Once every revolution the wheel closes a contact in series with a relay, thus providing the required seconds' intervals.

The regularity of these intervals is apt to be marred by (*a*) hunting of the motor, which may produce errors of a quarter of a millisecond or more[2], and (*b*) lack of exact repetition in the closing of the contact by the wheel. For example, if an accuracy of repetition of 10 microseconds is aimed at, and the wheel is 60 mm. in diameter, i.e. about 200 mm. in circumference, the tolerance in the making of the contact is only two thousandths of a millimetre, a requirement which few rotating contacts will satisfy.

There is, however, a purely electrical method of obtaining seconds' intervals[3], with an accuracy of repetition of 10 microseconds or better. It is well known that the variation of anode potential of a multi-vibrator is made up of an extremely

sharp change followed by an exponential return to the original condition. In fact the curve of potential is identical with that of the current in a condenser discharged through a resistance : if inductance were completely eliminated from the circuit the rise of current would be instantaneous. Now a multi-vibrator contains no inductance apart from that inevitably associated with valves, condensers and resistances. If, as in the apparatus described below, the values of condensers and resistances are fixed, and the latter are of the moulded type, the residual inductance is very small, as is also the coefficient L/CR^2, especially for the multi-vibrators of long period considered here.

The free period of a multi-vibrator depends primarily on the grid resistances and the grid-anode capacitances, and to a lesser extent on the anode resistances and the characteristics of the valves, but a multi-vibrator is readily dragged into step by a source of alternating current of period approximately equal to its free period. Control is effected by injecting into the multi-vibrator circuit a small e.m.f. from the controlling source, and it can be effected also by a source with a period which is approximately a multiple or sub-multiple of the free period of the multi-vibrator.

The production of seconds' intervals from the 1 kc./sec. source was effected in three steps by multi-vibrators having free periods of approximately 10 msec., 100 msec. and 1 sec. Instead of the two triode valves usually employed for each multi-vibrator stage, a single triode-pentode valve was used and the controlling voltage was injected on the suppressor grid of the pentode section. This method of injection was found to be far superior to the more usual methods of injection into anode or grid circuits. Resistance-capacity coupling between the triode anode of the preceding stage and the suppressor grid of the controlled stage was used, a capacity of 50 $\mu\mu$F. between the electrodes being found to be adequate when the suppressor grid was earthed via a fixed 1000-ohm resistance. This resistance was increased to a value of 10 000 ohms in the first stage, so that approximately 1·0 v. was impressed on the suppressor grid when a 10-v. signal at 1 kc./sec. was applied from the controlling source in series with 1 megohm.

All the components were fixed, the grid resistances being adjusted by trial. Ordinary half-watt moulded resistors with a manufacturing tolerance of ± 15 per cent vary sufficiently between themselves for the desired value to be selected from a dozen or so resistors. Paper condensers were used for the large capacitances.

The high-tension supply was nominally 72 v., and the total current taken by the three valves together was about 4 ma.

Adjustment of the free period was facilitated by making use of the fact that it decreases when the high tension is decreased. In the case of the first stage it is best to set the high tension at the lowest value which it is likely to have in practice, in the present case, say, 67 v., and to adjust the resistances in the grid circuits so that the free period only just exceeds the desired period, an adjustment which can be effected very conveniently with the help of the checking circuit

described below. The output from the 1 kc./sec. source being fixed at some convenient value, the resistance through which it is applied to the suppressor is altered so that the multi-vibrator is controlled even at the highest voltage at which it is desired to work, say 75 v. in the present case. It is then necessary to verify that control is still satisfactory at the lowest voltage, because naturally control cannot be effected beyond a certain voltage band, and the input necessary for control at, say, 78 v. might prove too large for control at 67 v. For the other two steps it is sufficient to adjust the grid resistances at the smallest value for which the stage remains under control at the lowest voltage. It was found that the voltage could then be raised by 10 per cent before control ceased.

Since the free period of a multi-vibrator increases when the filament voltage is increased, and decreases when the high tension is decreased, there is some measure of self-compensation for the changes in free period due to ageing of the batteries.

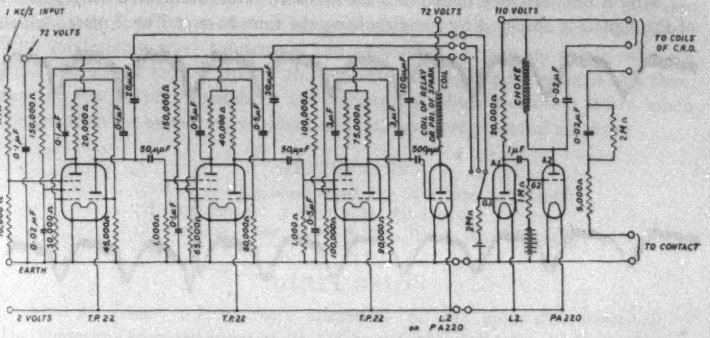


Figure 1. Multi-vibrators and checking circuit.

Control could also be effected from a 200-v. battery by means of a neon-lamp stabilizer working at 72 v., but the grid resistances had to be readjusted to suit the altered conditions.

The circuit used for ascertaining that the multi-vibrators are under control is in practice separate from the multi-vibrator panel, although shown here on the same circuit diagram, figure 1, for convenience. It is designed to register the sudden drops in potential of the anodes of the multi-vibrators, transmitted for the purpose through the 20, 30, and 100 $\mu\mu$F. condensers which can be connected in turn to the grid G 1. A sudden drop in potential of a triode anode thus causes a sudden reduction in current in the valve, with a sudden rise in potential of the anode A 1 and therefore of the negatively biased grid G 2. The valve A 2 G 2 in consequence passes current, direct current going through the choke and rapid variations through the condenser and the coils of the cathode-ray oscillograph, thus producing a signal which is superimposed on the time base. The time base of the oscillograph is synchronized by the 1 kc./sec. source, and, to verify

synchronism as well as to determine which multiple of 1 msec. the time base is synchronized with, a small 1 kc./sec. voltage from the source is applied to the "Y" plates, so as to produce a wavy line like the upper one in figure 2. Thus if the time base is adjusted for 10 waves and if the first multi-vibrator gives a stationary signal, as in the lower photograph in figure 2, it is being controlled at a frequency of 100 c./sec. With 9 or 11 waves the multi-vibrator signal would wander about from wave to wave.

In the same way the second multi-vibrator, if controlled, will give a signal every 100 msec., and this signal will be repeated at the same spot on the screen if the time base is synchronized to 10 msec. or 4 msec., both of which are factors of 100. Finally the third multi-vibrator gives a signal every second, and this signal occurs at the same spot if the time base is synchronized to 8 msec. or 10 msec., but not to 9 msec. or 11 msec.

After it has been ascertained that the multi-vibrators are in step the regularity of the signals is observed by synchronizing the time base at 1 or 2 msec., which

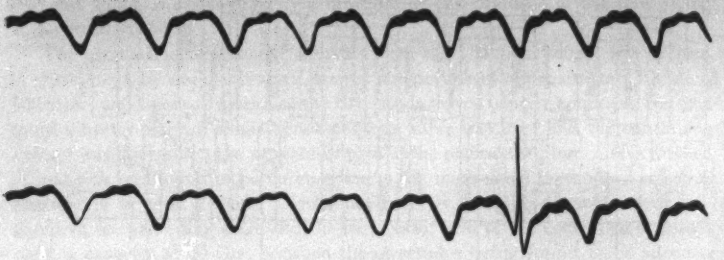


Figure 2.

extends the time scale to about 1 mm. for 25 μsec., thus rendering easily detectable any small variations of the time at which the condenser C 1 begins to charge. The result of the search for variations was negative; in other words the signals were found to be regular to at least 10 μsec.

The quick rise of current through condenser C 1 provides a signal every second, but signals can be obtained by other means if desired. For instance, the output of the last multi-vibrator, figure 1, is taken to the grid of a triode valve without a grid leak. The result is that during the negative part of the cycle the grid is blocked and there is no anode current, whereas during the positive part the normal anode current flows. Also, again on account of the rapidity of the fall of potential at the beginning of the negative part of the cycle, the fall of anode current of the triode is extremely sharp, and a good spark is obtained across the secondary of a high-ratio transformer inserted in the anode circuit. By this means the seconds can be recorded on a spark chronograph. The behaviour of a polarized relay inserted in the anode circuit was also examined by observing the time of closing of the contact after the anode current dropped to zero, when the conclusion was

reached that variations of the time delay of polarized relays of the type used do not exceed 10 μsec. Relays of that kind should, therefore, be suitable for recording calibration marks on high-speed chronographs, and of course for driving slave clocks provided with second, minute and hour hands.

Comparisons were made between the multi-vibrator outfit described above and two synchronous motors, one of which is probably as good as any other in the country. For the purpose the contact operated by the motor was made to charge condenser C 2, whilst in an alternative method the contact operated the relay, which in turn charged C 2. One wheel showed irregularities which sometimes reached 1/4 msec., whereas with the other wheel they were under 0·1 msec. Moreover the actual moment of make of the contact was not so well defined as the beginning of rise of potential in the multi-vibrators, and the charging curve did not assume the same shape each time.

The general conclusion is that the multi-vibrator outfit provides intervals of regularity superior to that of intervals obtained from a synchronous motor. With moderately good condensers and resistances there is no reason why the outfit should not keep in step with the controlling source for months at a time, and the few valves, resistances and condensers which constitute it are much cheaper than the synchronous motor and its associated powerful amplifier.

REFERENCES

(1) Essen. *Occasional Notes Roy. Astr. Soc.* **1**, 5 (1938).
(2) Tomlinson. *J. Brit. Astr. Ass.* (1939).
(3) George. *Bur. Stand. J. Res., Wash.*, **21**, 367 (1938).

DISCUSSION

Mr. L. Essen. I am very interested in the apparatus described by Dr. Vigoureux from the point of view of its possible application to the production of continuous seconds' impulses from frequency standards. At present it is usual to divide the frequency of the standard to 1000 c./sec. by multi-vibrators and to effect the further division to 1 c./sec. by synchronous motors. It is my experience that these motors operate continuously for about four months and then often need some attention such as cleaning and regreasing. If the multi-vibrator circuits described could be made to give a better or even equal reliability they would be preferable to the motor on account of the lower initial cost and lower power consumption. In 1934, when I was designing an additional standard, I considered this possibility and made an experimental frequency divider of this kind. Although the control at each stage was very strong and satisfactory operation was obtained for periods of 12 hours, occasional sudden changes occurred in the phase of the impulse. These were attributed at the time either to large disturbances in the mains supply voltages or to external electrical disturbances. In a time comparison it is usual to record say 30 impulses once every 24 hours, so that there is no means of detecting a change of phase such as that described. On the other hand, the synchronous motor would stop under such conditions, and although

the record would be lost a false result would not be obtained. In view of these results it was decided that the motor was more reliable, and although the multi-vibrator dividing circuit was shown at the National Physical Laboratory as an interesting exhibit it was not used as part of the standards equipment. In the model now described by Dr. Vigoureux and Mr. Stoakes more modern valves are used, enabling a much neater and more convenient assembly to be made, and I shall be pleased if he can give any information concerning its reliability over long periods.

For this particular application the shortness of the impulses is not usually of importance, since the impulses from standard pendulum clocks or time signals are considerably less precise than those obtained from a phonic motor; and if the comparison is between two frequency standards, very high accuracy indeed is obtained by measuring the beats between them. The sharp impulse would, however, be useful for comparisons between two frequency standards whose frequencies do not bear a simple ratio and between which it would therefore be difficult to obtain beats.

The sharpness of impulse can be used in other ways. A few years ago, for instance, I needed a very accurate measurement of the beat of 1 or 2 per second between two oscillators. One method I used was to control a multi-vibrator by the sinusoidal variations of voltage at the beat frequency, and thus to obtain sharp impulses at the same frequency. My accuracy was limited from other considerations to 0·0001 second, so I did not achieve the full precision of a few microseconds obtained by the authors.

AUTHORS. The sudden changes of phase referred to by Mr. Essen never occurred in our tests, which, however, have not covered periods of more than a day.

Lack of provision for automatic indication that a multi-vibrator has fallen out of step would admittedly be a weakness, but it is possible to arrange for an indicator to perform this operation. For instance, two independent outfits of five multi-vibrators each can be controlled from the same 100 kc./sec. crystal, and the signals from the last stages of the two chains applied respectively in opposite directions to the two coils of a relay. As long as the seconds' impulses occur simultaneously their effects oppose each other and the relay remains open, but if one of the multi-vibrators falls out of step the impulses occur at different times, the relay is closed, and operates a gas-discharge tube which works an alarm.

Our paper was being published when we became acquainted with the work of Uffelmann on the scale-of-two counter (see *Proc. Phys. Soc.* **51**, 1028 (1939)), which, although labouring under the double disadvantage of requiring a large number of stages and of yielding frequencies none of which is a decimal sub-multiple of the controlling frequency, yet seems more positive in action than either the synchronous motor or the multi-vibrator, and might be successfully adapted to time measurement in observatories.

THE USE OF UNIFORM CHROMATICITY SCALES

By J. G. HOLMES, A.R.C.S., B.Sc., D.I.C., F.Inst.P.,
Chance Brothers and Co. Ltd.

Received 22 *December* 1939

ABSTRACT. The C.I.E. trichromatic system is capable of modification to a similar system of three coordinates in which a diagram of colours has the property that the distance between the points representing two colours is approximately proportional to the subjective colour-difference.

It is shown that a similar effect can be obtained by a distortion of the C.I.E. co-ordinate lattice giving either the *Uniform Chromaticity System* proposed by Judd or the *Rectangular Uniform Chromaticity System* proposed by Breckenridge and Schaub. Nomographic transformations between the systems are also described.

The object of these transformations is to enable results which are recorded on the C.I.E. basis to be plotted or visualized on the U.C.S. system.

§1. INTRODUCTION

The trichromatic system of measuring and recording colour is well established and is now usually employed on the basis laid down by the Commission Internationale de l'Éclairage in 1931[1],[2],[3]. Suggestions have recently been made that the system would be of greater practical use and would be more easily understood if it were modified in such a way that the colour chart was a truer picture of the apparent disposition of colour. The C.I.E. system has the disadvantage that the distance between two points on the chart bears very little resemblance to the subjective difference between the two colours which they represent, with the result that, for example, a wide range of orange colours is given by a small area whilst a smaller range of green colours extends over a large area of the chart.

The possibility of overcoming this and similar apparent anomalies has been considered in several ways by Smith[4] and MacAdam[8]. The mathematical expressions on which the system is based contain eight arbitrary constants which may be used in order to obtain different dispositions of the coordinates without losing the fundamental characteristics of the standard observer which form the basis of the C.I.E. system.

A coordinate system was found by Judd[5],[6] which has the property that the length of any line on the chart is an approximate measure of the chromaticity difference between the two colours represented by the end points of the line. It is not physically possible to derive a chart for which this requirement is critically fulfilled even for one standard condition of observation, and, if it were possible, different charts would be needed for different conditions[7]. There is also a lack

of experimental data on which the system may be based, with the result that a number of different systems could be used all of which are within the experimental error of the published observations—proposals have been put forward and discussed by MacAdam[8], Sinden[9] and Breckenridge and Schaub[10].

There is, however, a use for an approximate chart provided its approximate nature is never forgotten. This note is intended to give a simple statement of the relation between the uniform chromaticity system (U.C.S.) and the C.I.E. system, and of the ways in which the U.C.S. chart may be used.

§ 2. THE C.I.E. SYSTEM

Three primary stimuli, XYZ, are defined, of which X is analogous to red, Y to green and Z to blue. A colour may be expressed in terms of the amounts of each necessary to match it, in the form:

$$C=x'X+y'Y+z'Z.$$

The stimulus Y is the only one possessing luminosity, so that the luminosity of the colour C is given by y', the coefficient of Y. For most purposes the absolute value of the luminosity is immaterial, and we can write:

$$C=xX+yY+zZ,$$

where

$$x+y+z=1.$$

The colour is then precisely defined by the values of the coefficients x and y only, and all colours may be plotted on a chart such as figure 1, in which the Y-coefficient is the ordinate and the X-coefficient is the abscissa. This chart gives the spectrum locus and also the curves of equal colorimetric saturation[11]. It will be noticed that the saturation of purple colours has been calculated by the empirical "spectrum locus" joining the red and blue in accordance with the recent recommendation of the C.I.E.[12]. This practice gives positive figures for the saturation and includes all possible natural colours. This chart may be used for diagrammatic representation of colour in any way desirable, such as for definition of tolerances[13] or for the representation of gradual colour change[14], but is subject to the disadvantages stated above.

§ 3. THE U.C.S. SYSTEM

The modification proposed by Judd is a Maxwell triangle derived from the C.I.E. definitions by linear transformations. The following relations give the same results when plotted on rectangular coordinates:

$$r'=\frac{-0{\cdot}2028x+8{\cdot}6441y+0{\cdot}8529}{-1{\cdot}2028x+7{\cdot}6441y+1{\cdot}8529},$$

$$g'=\frac{-3{\cdot}9726x+2{\cdot}0086y+0{\cdot}6580}{-1{\cdot}2028x+7{\cdot}6441y+1{\cdot}8529},$$

$$b'=1-r'-g',$$

where r', g' and b' are the trichromatic coefficients of the colour.

This gives the result shown in figure 2, where the spectrum locus and the Planckian locus are plotted on $R'-G'$ rectangular coordinates. This is a

"Uniform Chromaticity Scale", and it will be seen that it is very different from figure 1. The red-orange-yellow range is extended whilst the white-yellow distance is reduced and the area of green colours is considerably reduced.

The same result may be obtained without the labour of the above transformation from the XYZ basis. As the transformation is linear or, in other

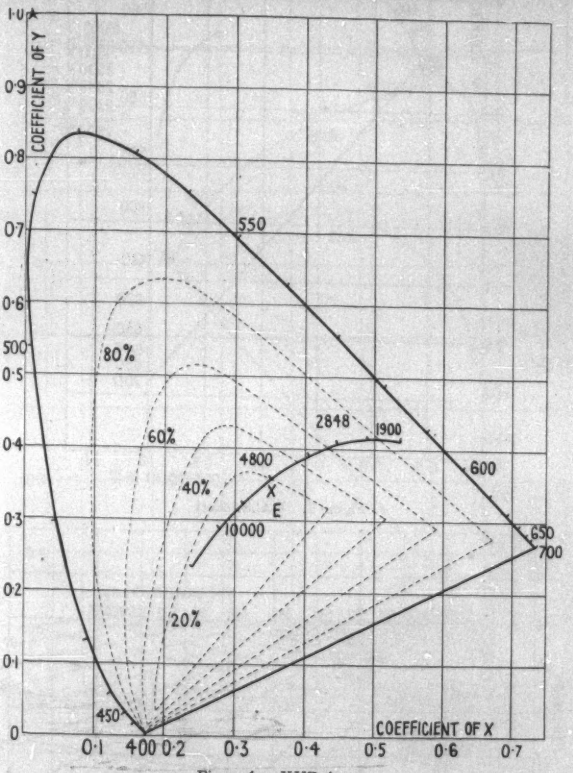


Figure 1. XYZ chart.

words, as it is a projection from one plane to another, the points which are collinear on figure 1 will also be collinear on figure 2 and the ordinates and abscissæ of figure 1 can be plotted on figure 2. This has been done in figure 3, which gives the same curve as Judd, but can be used to illustrate XYZ data on a basis of approximately uniform chromaticity.

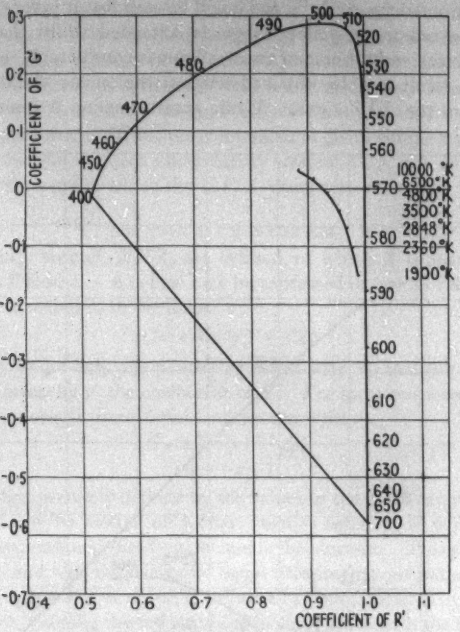


Figure 2. U.C.S. chart

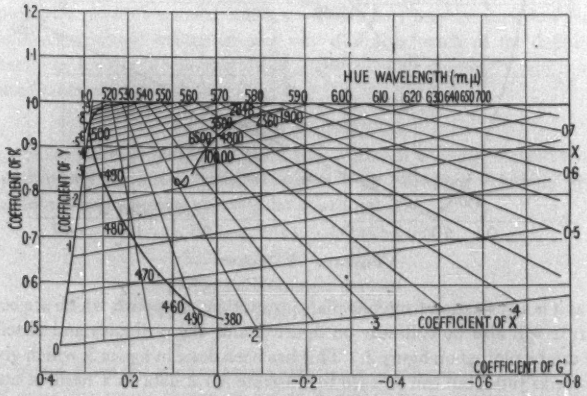


Figure 3. *XYZ* chart (U.C.S.).

§ 4. THE R.U.C.S. SYSTEM

This system was proposed by Breckenridge and Schaub[10] and is a further simplification of Judd's U.C.S. system. It is called the "Rectangular Uniform

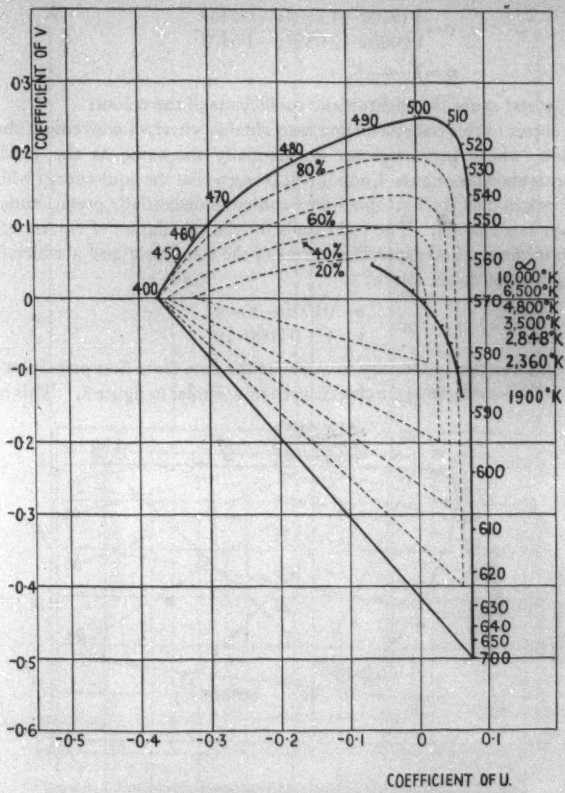


Figure 4. R.U.C.S chart.

Chromaticity Scale" (R.U.C.S.) and seems to bring the ideal of a uniform chart more into line with the requirements of industry and general practice. A series of equations was derived which gave a result within the margin of error

of the experimental data reported by Judd, and the transformation equations may be re-written as follows:

$$u=\frac{-0{\cdot}7480x-1{\cdot}3520y+0{\cdot}7000}{1{\cdot}000x-7{\cdot}0534y-1{\cdot}6402},$$

$$v=\frac{3{\cdot}1970x-1{\cdot}5504y-0{\cdot}5487}{1{\cdot}0000x-7{\cdot}0534y-1{\cdot}6402},$$

$$w=1-u-v,$$

where u, v and w are the trichromatic coefficients of the colour.

The object of this transformation is to obtain a chart, of convenient shape and orientation, whose properties are substantially the same as the Judd chart. The chart is shown in figure 4, and it will be seen that the equi-energy white point is at the origin and the four quadrants contain substantially green, blue, purple and red-yellow colours. The full details for the calculation of colorimetric data are given in the original paper[10] in terms of the symbols x' and y' which bear the following relation to the above:

$$u=0{\cdot}0750-x'=x'',$$
$$v=y'-0{\cdot}5000=y''.$$

It is not, however, necessary to calculate the data from first principles or even from the XYZ-coefficients if a chart is prepared similar to figure 3. This has been

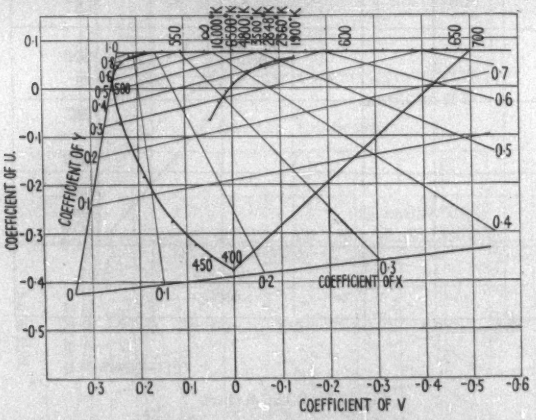


Figure 5. XYZ chart (R.U.C.S.).

done in figure 5, where it is possible to plot XYZ-coefficients directly on the R.U.C.S. chart. The coordinates from which figure 5 was plotted are given in Appendix I. A number of large copies of this diagram have been prepared.

The transformation from XYZ-coefficients to the UVW-coefficients may be rapidly made by means of a nomogram[11], [15]. The accuracy is not greater than about 0·003, but it is within the limits of accuracy of the uniform chromaticity

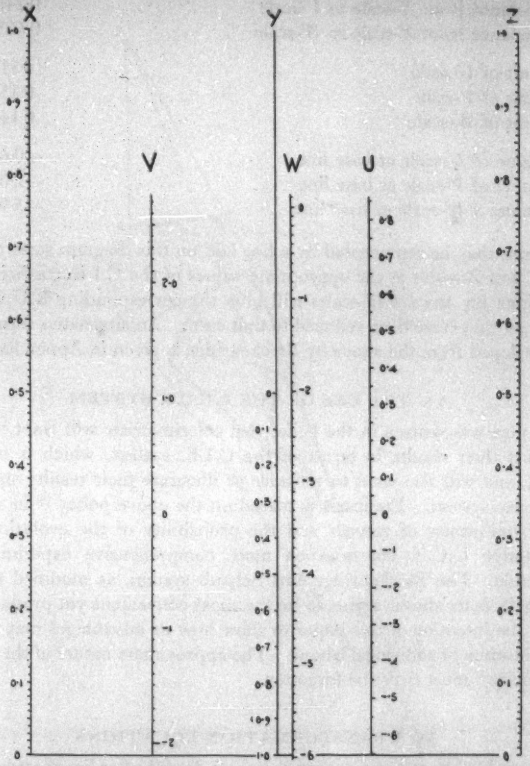


Figure 6. Nomogram for conversion of XYZ data to R.U.C.S. data.

basis and is probably close enough for any purpose for which the R.U.C.S. chart is likely to be used at present. A copy of the nomogram is given in figure 6, which is drawn from the following data:

Distance from X-scale to Y-scale and from Y-scale to Z-scale	$D/2$
Unit of X-scale and Z-scale	d
Unit of Y-scale	$-d/2$

Value of X-scale and Z-scale at base line	0
Value of Y-scale at base line	+1
Distance from X-scale to U-scale	0·6912D
Distance from X-scale to V-scale	0·2462D
Distance from X-scale to W-scale	0·6057D
Unit of U-scale	0·5112d
Unit of V-scale	0·1588d
Unit of W-scale	0·1459d
Value of U-scale at base line	−0·6520
Value of V-scale at base line	−2·0991
Value of W-scale at base line	−5·9425

A colour may be represented by a line laid on this diagram so as to intersect the X, Y and Z-scales at the appropriate values of the C.I.E. coefficients. The intersections on the UVW-scales will give the corresponding R.U.C.S. coefficients when they have been reduced to unit sum. An alternative form of nomogram developed from the above by Breckenridge is given in Appendix II.

§ 5. THE USE OF THE R.U.C.S. SYSTEM

This note was written in the belief that colorimetrists will want to calculate and record their results in terms of the C.I.E. system, which is now almost universal, and will also want to visualize or illustrate their results on a uniform chromaticity system. Emphasis is placed on the above policy [8] in view of the need for uniformity of records and the probability of the evolution of more representative U.C.S. charts when more comprehensive experimental data are available. The Breckenridge and Schaub system, as modified to give the UVW-coefficients above, seems to be the most convenient yet produced, and it has been the intention of this paper to show how its advantages may be enjoyed with a minimum of additional labour. The approximate nature of the "Uniform Chromaticity" must never be forgotten.

§ 6. TRANSFORMATION EQUATIONS

The two U.C.S. systems quoted above are described as linear transformations of the C.I.E. system and consist of changes of shape and disposition of the C.I.E. colour chart, in which straight lines remain straight. The mathematical process of this change is fully described in the references cited, but, in the words of MacAdam [8], "A familiarity with the basic theories of projective geometry brings the behaviour of transformations of the chromaticity diagram within the scope of intuitive reasoning". A simple experiment to illustrate this is shown diagrammatically in figure 7. A very similar diagram has been given by Sinden [9] in his discussion of mechanical-optical transformations. A transparency of the

C.I.E. colour chart is prepared and is shown as the horizontal plane *XYZ* in figure 7, with a point source of light above it at P. The rays of light transmitted, which form the projection of the transparency, are received on an oblique plane below it, so that by adjustment of the position of the source and the orientation of

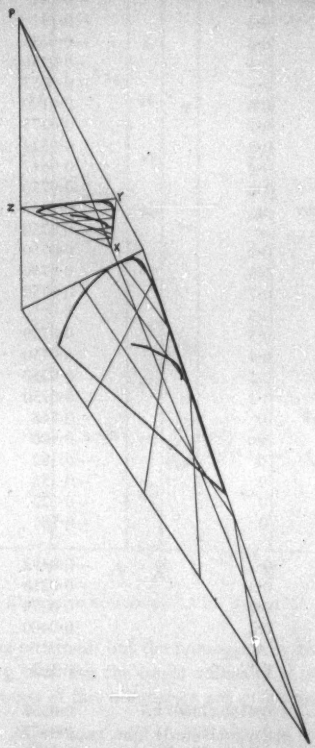


Figure 7. Diagram showing linear projection of charts from *XYZ* to U.C.S.

the lower plane any linear transformation of the transparency may be obtained. The assembly may be inverted by using a lens of small aperture in the place of the light source and reversing the direction of the light through the transparency, which gives better definition where this experiment is used for purposes of accurate demonstration.

Appendix I

Coordinates for plotting X and Y on R.U.C.S. chart

X	Y	u	v
0	0	−0·425	0·334
0	0·1	−0·241	0·300
0	0·2	−0·141	0·282
0	0·3	−0·0785	0·270
0	0·4	−0·0356	0·261
0	0·5	−0·0053	0·256
0	0·6	0·0190	0·252
0	0·7	0·0375	0·248
0	0·8	0·0524	0·245
0	0·9	0·0647	0·243
0	1·0	0·0750	0·241
0	∞	0·1915	0·220
−∞	∞	0·0750	0·591
0·1	0·9	0·0750	0·205
0·2	0·8	0·0750	0·162
0·3	0·7	0·0750	0·108
0·4	0·6	0·0750	0·038
0·5	0·5	0·0750	−0·059
0·6	0·4	0·0750	−0·194
0·7	0·3	0·0750	−0·400
0·8	0·2	0·0750	−0·754
−∞	0	−0·748	3·197
0·1	0	−0·406	0·148
0·2	0	−0·382	−0·063
0·3	0	−0·355	−0·306
0·4	0	−0·323	−0·588
0·5	0	−0·286	−0·920
0·4	0·2	−0·0492	−0·158
0·5	0·2	−0·0218	−0·290
0·6	0·2	0·0079	−0·432
0·7	0·2	0·0400	−0·587
0·5	0·1	−0·103	−0·485
0·4894	0·4150 (2360° K.)	0·0557	−0·0912
0·4476	0·4074 (Illt. A)	0·0457	−0·0616
0·3484	0·3516 (Illt. B)	0·0096	−0·0053
0·3100	0·3162 (Illt. C)	−0·0113	0·0134
1/3	1/3 (Illt. E)	0·000	0·000
0·7347	0·2653 (700 mμ.)	0·075	−0·500
0·0105	0·5134 (499 mμ.)	0·000	0·250
0·1741	0·0050 (380 mμ.)	−0·375	0·000

Appendix II

In correspondence with F. C. Breckenridge, a modified form of nomogram was evolved as shown in figure 8. The arrangement of the variables is such that

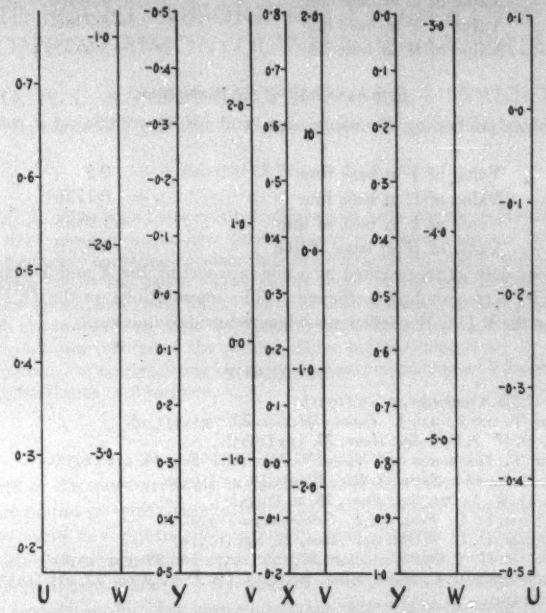


Figure 8. Alternative nomogram. *XYZ* data to R.U.C.S data.

an improved accuracy is obtained, but the nomogram is divided into two halves, the left-hand side being used for the larger values of Z and the right-hand for the smaller. The constants of this nomogram are as follows:

Right-hand half of the nomogram

Distance from X-scale to Y-scale	D
Unit of X-scale	d
Unit of Y-scale	$-d$
Value of X at base line	$-0{\cdot}2$
Value of Y at base line	$1{\cdot}0$
Distance from X-scale to U-scale	$2{\cdot}2385D$
Distance from X-scale to V-scale	$0{\cdot}3266D$
Distance from X-scale to W-scale	$1{\cdot}5363D$

Unit of U-scale	$1{\cdot}6556d$
Unit of V-scale	$0{\cdot}2106d$
Unit of W-scale	$0{\cdot}3701d$
Value of U at base line	$-0{\cdot}5024$
Value of V at base line	$-2{\cdot}7387$
Value of W at base line	$-5{\cdot}6525$

Left-hand half of the nomogram

All values containing D change sign, and values containing d remain unchanged.

Value of Y at base line	$0{\cdot}5$
Value of U at base line	$0{\cdot}1736$
Value of V at base of line	$-1{\cdot}9635$
Value of W at base of line	$-3{\cdot}5770$

A colour may be represented by a line intersecting the X and Y-scales at the values of its corresponding coefficients. The intersections on the U, V and W-scales give the R.U.C.S. coefficients when reduced to unit sum.

REFERENCES

(1) *Proc. C.I.E. Cambridge*, p. 19 (1931).
(2) SMITH, T. and GUILD, J. *Trans. Opt. Soc.* **33**, 73 (1931–2).
(3) JUDD, D. B. *J. Opt. Soc. Amer.* **23**, 359 (1933).
(4) SMITH, T. Discussion on " Vision ", *Proc. Phys. Soc.* **44**, 212 (1932).
(5) JUDD, D. B. *Bur. Stand. J. Res., Wash.*, **14**, 41 (1935).
(6) JUDD, D. B. *J. Opt. Soc. Amer.* **25**, 24 (1935).
(7) MACADAM, D. L. *Proc. C.I.E., Scheveningen* (1939).
(8) MACADAM, D. L. *J. Opt. Soc. Amer.* **27**, 294 (1937).
(9) SINDEN, R. H. *J. Opt. Soc. Amer.* **27**, 124 (1937), and **28**, 339 (1938).
(10) BRECKENRIDGE, F. C. and SCHAUB, W. R. *J. Opt. Soc. Amer.* **27**, 225 (1937), and **29**, 370 (1939).
(11) HOLMES, J. G. *Proc. Phys. Soc.* **47**, 400 (1935).
(12) *Proc. C.I.E., Scheveningen* (1939).
(13) BSS. no. 563 Aviation Colours.
AAR 69–38 Railway Signal Colours.
(14) MCNICHOLAS. *Bur. Stand. J. Res., Wash.*, **17**, 935 (1936);
HOLMES, J. G. *Proc. Lighthouse Conf. Berlin* (1937).
(15) WINCH, G. T. *Trans. Illum. Engng Soc., Lond.*, **2**, 137 (1937).

THE SOUND-ABSORBING PROPERTIES OF SOME COMMON OUT-DOOR MATERIALS

By G. W. C. KAYE, O.B.E., F.R.S., and E. J. EVANS, B.Sc.,
Physics Department, The National Physical Laboratory, Teddington, Middlesex

Received 25 *January* 1940

ABSTRACT. The sound absorption coefficients of pure sounds of frequencies from 125 to 4000 cycles per second have been measured at random incidence for specimens of some commonly occurring out-door materials, among them, gravel, turf, sand, ashes, railway-track ballast and snow. They mostly share the common characteristic of increased absorption with rising pitch. Some of the materials (e.g. loose gravel soil and ashes) are highly absorbent, snow is remarkably so, while others (e.g. compressed gravel and wet sand) are indifferent absorbents. Some practical aspects are discussed, and also the influence of nearly grazing incidence, such as often obtains with out-door sounds, in raising the degree of absorption.

§ 1. INTRODUCTION

SOME of the more arresting acoustic phenomena encountered out of doors are bound up with the reflection of the high-pitched components of noises by every-day objects such as fences, trees, hedges, telegraph poles, etc., or the selective absorption of such components by the air and other bodies, for example, in distant echoes. Such matters are of general interest and experience; and the present enquiry has been undertaken with a view to obtaining wider and more specific information on the acoustic properties of some of the surfaces commonly met with in the open air, for which such data are not available, so far as we are aware. Measurements of the acoustic absorption coefficients of some typical natural surfaces, such as turf, sand, gravel soil, snow [1], and other materials of related interest have accordingly been made by the reverberation method at a variety of pitches. The information may be of significance in connection with acoustical measurements made out of doors, where the nature of the bounding surfaces would be expected to exercise influence, for example, in such matters as the audibility of speech in the open air.

§ 2. MEASUREMENTS

The measurements followed the normal procedure used in determining sound-absorption coefficients by the reverberation method at the National Physical Laboratory [2]. The reverberation chamber is irregular in shape, with non-parallel walls and a sloping ceiling. The double walls are mainly of rendered

brickwork, while the floor is of rendered concrete, the various surfaces being painted. The volume of the chamber is about 10,000 cubic feet and the total surface area about 3000 square feet. The acoustic isolation of the chamber is unusually high—about 90 db. A check is kept on the air temperature and humidity: in the present tests they ranged from 13° to 20° C., and 40 to 80 per cent respectively.

For the tests instrumental measurements were made of the decay of reverberant sound in the chamber, both when empty and when the various materials were present. The source of sound was a loudspeaker, and the test frequencies were a series of pure notes from 125 to 4000 cycles per second, that is, covering virtually the speech range. To reduce the effects of resonance and interference phenomena in the chamber, notes of fluctuating or "warbling" pitch were used, the frequency varying five times a second between limits of ± 2 to ± 10 per cent (depending on the frequency) from the mean. In the earlier experiments, the rates of decay were measured by means of a timing device consisting of a thyratron relay and clock, in conjunction with a microphone and amplifier [2]. More recently, graphical records of the decay were obtained by means of a sound-level recorder of the type described by Braunmühl and Weber [3].

The mean sound absorption coefficients (i.e. the fractions of the incident sound-energy absorbed by the specimen under the conditions of test) were calculated on the basis of the traditional Sabine formula, as modified by Eyring, Schuster and Waetzman, and Knudsen, namely :—

$$T = \frac{1}{20} \frac{V}{-S.\log_e(1-a)+4mV},$$

where T is the standard reverberation period of the chamber, i.e., the time taken for the intensity of the reverberant sound to decay a millionfold, or 60 decibels, V is the volume of the chamber in cubic feet, S is the total area in square feet of the bounding surfaces of the chamber, a is the average sound-absorption coefficient of the bounding surfaces of the chamber, i.e., $a = \Sigma a_1 s_1 / S$, where s_1 is the area of each material constituting the bounding surface, and a_1 is its absorption coefficient, whilst m is the air-attenuation coefficient, which operates appreciably for frequencies greater than about 2000 cycles per second.

The decay measurements were mostly made over an intensity drop of about 60 decibels. For frequencies of 500 cycles per second and above, the average decay was usually uniformly logarithmic. At lower frequencies, however, when the wave-length became comparable with the dimensions of the specimen and the room, the decay was less uniform, restricting somewhat the accuracy of the results, especially at the lowest frequency of 125 cycles per second.

The test materials were placed on the floor of the chamber, to form in each case a single specimen 10 feet by 10 feet in area, in accordance with the size adopted as standard in reverberation tests at the Laboratory. In the case of most of the materials, a hollow wooden frame 12 inches deep was used to enclose the specimens of various thickness. Figure 1 shows a photograph of a specimen

so enclosed, consisting of rough turf laid on compressed gravel. In the case of the snow, this was deposited in the open on thin plywood boards in sections about 2 feet 6 inches by 2 feet laid on the ground. The sections were then assembled on the floor of the chamber, the surface of the snow being disturbed as little as possible. The gravel subsoil was of local origin, excavated when the

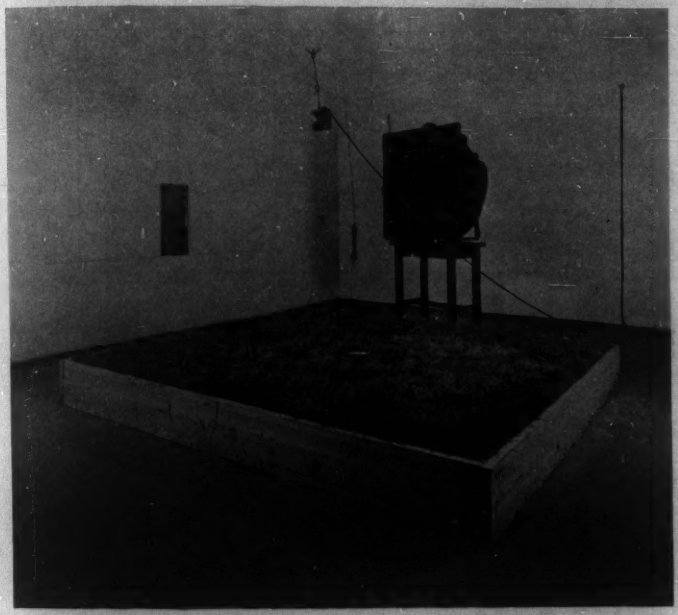


Figure 1. Test specimen of turf on gravel in reverberation chamber.

ground was in a damp condition. The experiments on sand were conducted with ordinary builders' sharp sand, which was first thoroughly dried by prolonged heating. The granite, limestone and slag were materials used as ballast on railway tracks, while the ashes came from the furnace of a railway locomotive. The foamed slag consisted of small porous stones, and is, we undertand, a by-product in the manufacture of pig iron in blast furnaces.

§ 3. RESULTS

The values of the absorption coefficients obtained are given (to the nearest 0·05) in the following table. Figure 2 displays the results for materials about 12 inches thick, while figure 3 illustrates the variation of absorption with thickness in the cases of ashes and snow.

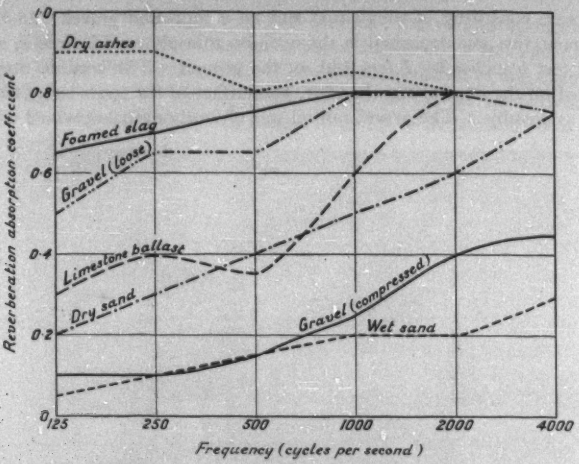


Figure 2. Reverberation absorption coefficients for materials about 12 inches thick.

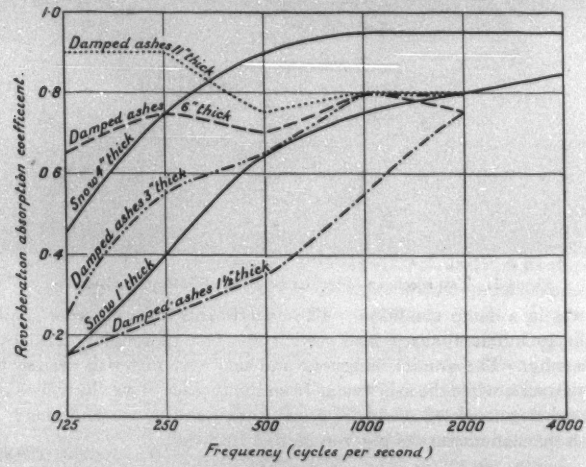


Figure 3. Variation of absorption with thickness in the case of damped loose ashes and snow.

The results show a number of points of interest. The increase of absorption with rising pitch is a characteristic shared by nearly all the materials. To take first the materials of the full thickness of about 12 inches (figure 2), it is

Table

Material as tested (area of specimen 10 feet by 10 feet)	Thickness (inches)	Weight (lbs. per cubic foot)	Reverberation absorption coefficients at frequencies of					
			125 (c./s.)	250 (c./s.)	500 (c./s.)	1000 (c./s.)	2000 (c./s.)	4000 (c./s.)
Gravel soil, loose and moist	4	87	0·25	0·60	0·65	0·70	0·75	0·80
,, ,, ,, ,, ,,	8	,,	0·45	0·55	0·70	0·75	0·70	0·65
,, ,, ,, ,, ,,	12	,,	0·50	0·65	0·65	0·80	0·80	0·75
,, ,, moist, pounded to give firm surface	12	101	0·10	0·10	0·15	0·25	0·40	0·45
Rough turf on 10″ of compressed gravel	12–13	—	0·15	0·25	0·40	0·55	0·60	0·60
Sand (sharp), dry	4	108	0·15	0·35	0·40	0·50	0·55	0·80
,, ,, ,,	8	,,	0·15	0·30	0·45	0·50	0·55	0·75
,, ,, ,,	12	,,	0·20	0·30	0·40	0·50	0·60	0·75
,, wet (14 lbs. water per c. ft.)	4	122	0·05	0·05	0·05	0·05	0·05	0·15
,, ,, ,, ,, ,, ,, ,,	8	,,	0·05	0·05	0·10	0·10	0·10	0·15
,, ,, ,, ,, ,, ,, ,,	12	,,	0·05	0·10	0·15	0·20	0·20	0·30
Granite ballast (1500 stones per c. ft.)	11	86	0·20	0·45	0·25	0·50	0·60	—
Limestone ballast (1500 stones per c. ft.)	11	81	0·30	0·40	0·35	0·60	0·80	—
Slag ballast (900 stones per c. ft.)	11	81	0·30	0·45	0·35	0·65	0·90	—
Foamed slag, graded between $\frac{1}{4}$ inch and $\frac{1}{2}$ inch	11–12	26	0·65	0·70	0·75	0·80	0·80	0·80
Ashes, dry, loose (mainly fine cinders)	11	34	0·90	0·90	0·80	0·85	0·80	—
,, damped, loose (2·5 lbs. water per c. ft.)	11	36	0·90	0·90	0·75	0·80	0·80	—
,, ,, ,, ,, ,, ,,	6	,,	0·65	0·75	0·70	0·80	0·75	—
,, ,, ,, ,, ,, ,,	3	,,	0·25	0·55	0·65	0·80	0·80	—
,, ,, ,, ,, ,, ,,	$1\frac{1}{2}$	,,	0·15	0·25	0·35	0·55	0·75	—
,, ,, pounded and rolled	$9\frac{1}{2}$	42	0·75	0·70	0·75	0·80	0·80	—
Snow, freshly fallen	1	—	0·15	0·40	0·65	0·75	0·80	0·85
,, ,, ,,	4	—	0·45	0·75	0·90	0·95	0·95	0·95

evident that loose gravel, ashes and foamed slag are very good absorbents, that dry sand and the ballast materials are moderately absorbent, while compressed gravel and wet sand have comparatively low absorbing powers.

The absorbing power of the loose gravel was greatly decreased when it was compressed. The loose gravel was probably fairly representative of freshly dug ground, while the compressed gravel corresponded roughly to firm undisturbed soil. Ashes, on the other hand, did not show a marked reduction in efficiency when compressed by pounding and rolling with a 3 cwt. roller. It will be seen that turf laid on firm gravel soil is quite an effective absorbent, considerably better than the underlying gravel alone.

As regards the effect of moisture, in the case of ashes 11 inches thick, the water added by sprinkling, which was equivalent to about $\frac{1}{2}$ inch of rainfall, did not greatly affect the acoustical absorption. In the experiments on sand, however, enough water was added virtually to produce saturation, with the effect of greatly reducing the absorption. It was not practicable to obtain a uniform distribution of water throughout the sand, on account of the tendency of the water to gravitate through the material. The proportion of water near the surface was less as the thickness increased, and this is probably the reason for the increase of absorption with thickness in the case of the wet sand.

For loose gravel and dry sand, the effect of thickness appears to be small, indicating that sound does not penetrate very far into these materials. It is to be anticipated that the behaviour of the compressed gravel would be similar. As may be seen from figure 3, ashes and snow, however, show in an interesting way the effect of thickness in increasing the absorption, especially at low frequencies. In the case of the greater thicknesses of ashes, there appears to be some indication of a resonance effect in the region of 125 and 250 cycles per second.

The fresh snow is seen to be remarkably absorbent. In particular, a 4-inch layer absorbs as much as 90 per cent or more of the intensity of medium and high-pitched notes ; and greater thicknesses would no doubt present equally striking figures for low-pitched notes. Thus the pronounced muting effect of a layer of freshly fallen snow can be well understood. An interesting illustration of the high absorption of snow was described in a recent letter to *Nature* by Seligman[4], who relates how during the construction of a shaft in the snow on the Jungfrau it was impossible, when the depth had attained 35 feet, for those at the top to hear the voices of the men working below, even if raised in a shout.

§ 4. DISCUSSION

The present results show that some of the surfaces commonly encountered out of doors have high sound-absorption coefficients. On the other hand, water, hard ground and paved or concrete surfaces have absorption coefficients as small as 0·05, or even less. There is, therefore, an extensive range in the acoustic properties of ordinary out-door surfaces and, whilst a quantitative evaluation

would be difficult, it is of interest to examine the general bearing of the present results on the propagation of sound in the open air. Although the influence of atmospheric conditions such as wind, humidity and temperature gradients is well understood, little attention appears to have been given to the nature of the surface over which the sound is propagated. We note in this connection, a remark made by W. C. Sabine[5] over twenty years ago, in discussing the conditions introduced in acoustical measurements by the surrounding materials: " If measurements are made in the open air, over a lawn, as was done by the late Lord Rayleigh in certain experiments, is due consideration given to the fact that the surface has an absorbing power for sound of from 40 to 60 per cent? ". Incidentally, the values mentioned by Sabine for grass, which appear to be no more than an estimate on his part, are in agreement with the present results for rough turf for frequencies from 500 to 4000 cycles per second.

It is to be noted that sound absorption coefficients determined by the reverberation method, as in the present experiments, give an average coefficient for all angles of incidence, the sound being incident in a random manner on the material. Paris[6], and Morse[7] have considered theoretically the variation of sound absorption with angle of incidence, and experiments supporting the theory have been described by Kühl and Meyer[8], Willig[9], and Hunt[10]. It is clear that sound striking a surface perpendicularly is less absorbed than sound striking it obliquely. More exactly, it appears in the cases of the materials so far examined, that as the angle of incidence is increased, the absorption generally rises to a maximum at large angles of incidence and subsequently decreases. It is therefore probable that there is considerable absorption when sound falls on most ordinary types of absorbing surfaces at nearly grazing incidence, corresponding, say, to speech in the open air when both speaker and listener are near the ground. The reduction of intensity caused in this way, when sound passes over an absorbing surface, has been observed experimentally by Békésy[11].

The above considerations indicate that for positions near the ground, speaking conditions will depend on the nature of the surface as well as on atmospheric conditions. Thus it is to be expected that it would be easier to talk over water or concrete than over grass, whilst greater difficulty would be experienced over newly dug ground or freshly fallen snow. Knudsen[12] has described measurements on the range of audibility of speech which were carried out in a desert. From the similarity of the present results for dry sand and rough turf, it would seem that conditions in a desert correspond approximately to speaking over fairly short grass.

As regards practical applications, the experiments on ballast materials, which were undertaken for the London, Midland & Scottish Railway, indicate that proper choice of ballast materials may usefully contribute to noise reduction in trains, and that in this respect, ashes, though apt to be dust-producing, would possess advantages. The experiments on foamed slag were made on behalf of the London Underground Railway, and this material, provided with a suitable

perforated cover, has been successfully used for noise reduction purposes in tube tunnels. With reference to the results for turf, the practice (more common abroad than in this country) of growing grass between tram lines, is no doubt beneficial for reducing noise. The provision of lawns is similarly a useful factor in promoting quietness in areas surrounded by buildings, as, for example, when blocks of flats are erected around a central courtyard. Since the introductino of pneumatic tyres on road vehicles, there is little to choose acoustically between modern road surfaces, but nevertheless observant occupants of reasonably quiet cars find macadam roads more sound absorptive than concrete roads.

Some data for various outdoor surfaces have also been given by Eisner and Kruger[13], whose experiments were made in connection with acoustic reflection methods of determining the height of aircraft. The measurements were made from an airship, using a source of sound of frequency 2900 cycles per second, the absorption coefficients being measured for sound at normal incidence. Chief interest attaches to the value of 0·5 given for a field, which agrees with the present value of 0·6 for turf at 2000 and 4000 cycles per second, when allowance is made for the fact that absorption coefficients for sound at normal incidence are usually less than the corresponding reverberation values[14]. The corresponding figure which they obtained for a pine wood ranged between 0·2 and 0·4. They also mention that freshly ploughed soil would give high absorption, and this is supported by our results for loose gravel soil.

Finally, reference may be made to the sound-absorption figures for dried mud, in the form of rough bricks (with 2·5 per cent of straw by weight) given by Constable and Mahas[15]. Their reverberation results range from 0·32 at 250 cycles per second to 0·55 at 1000 and 2000 cycles per second. If the surface was smoothed, these figures were reduced to 1/3 of their former values.

We wish to thank the London, Midland & Scottish Railway Co. for permission to publish the acoustic absorption coefficients for railway-ballast-materials and ashes, and the London Passenger Transport Board for permission to include similar results for foamed slag. Mr. G. E. Ashwell and Mr. A. E. Bishop rendered valuable assistance with the measurements.

§ 5. SUMMARY

The sound absorption coefficients of pure sounds of frequencies from 125 to 4000 cycles per second have been measured at random incidence for specimens of some commonly occurring out-door materials, among them, gravel, turf, sand, ashes, railway-track ballast and snow. They mostly share the common characteristic of increased absorption with rising pitch. Some of the materials (e.g. loose gravel soil and ashes) are highly absorbent, snow is remarkably so, while others (e.g. compressed gravel and wet sand) are indifferent absorbents. Some practical aspects are discussed, and also the influence of nearly grazing incidence, such as often obtains with out-door sounds, in raising the degree of absorption.

REFERENCES

(1) KAYE, G. W. C. and EVANS, E. J. *Nature, Lond.*, **143**, 80 (1939).
(2) KAYE, G. W. C. *J. Acoust. Soc. Amer.* **7**, 167 (1936).
(3) BRAUNMÜHL, H. J. V. and WEBER, W. *Elekt. Nachr.-Tech.* **12**, 223 (1935).
(4) SELIGMAN, G. *Nature, Lond.*, **143**, 1071 (1939).
(5) SABINE, W. C. *Collected Papers on Acoustics*, p. 277 (1922).
(6) PARIS, E. T. *Proc. Roy. Soc.* A, **115**, 407 (1927).
(7) MORSE, P. M. *Vibration and Sound*, p. 303 (1936).
(8) KÜHL, V. and MEYER, E. *S.B. preuss. Akad. Wiss.* 416 (1932).
(9) WILLIG, F. J. *J. Acoust. Soc. Amer.* **10**, 293 (1939).
(10) HUNT, F. V. *J. Acoust. Soc. Amer.* **10**, 216 (1939).
(11) BÉKÉSY, G. V. *Z. tech. Phys.* **14**, 6 (1933).
(12) KNUDSEN, V. O. *Architectural Acoustics*, p. 497 (1932).
(13) EISNER, F. and KRUGER, K. *Hochfrequenztech. u. Electroakust.* **42**, 64 (1933).
(14) DAVIS, A. H. and EVANS, E. J. *Proc. Roy. Soc.* A, **127**, 89 (1930).
(15) CONSTABLE, F. H. and MAHAS, M. K. *Nature, Lond.*, **144**, 33 (1939).

Note added in proof

W. Ernsthausen and W. v. Wittern (*Akust. Z.* **4**, 353 (1939)) have recently described measurements of the acoustic reflecting powers of various types of ground, made in the open air for frequencies from 500 to 10,000 cycles per second. A loudspeaker suspended 20 metres above the ground (from a fire-escape ladder) emitted downward impulses consisting of short trains of sound waves, and relative measurements were made of the direct and reflected impulses by a microphone mounted vertically below the loudspeaker, some 2 metres above the ground. Allowing for the fact that the experiments were confined to normal incidence, their results for snow are in good agreement with ours, while there also appears to be fair agreement in the case of dry sand. For meadow land, however, results are given which vary over a wide range, the absorption coefficients being in some cases much less than would be expected from our results for turf.

THE BAND SPECTRUM OF SnTe IN EMISSION

By R. F. BARROW,
Departments of Inorganic Chemistry and Physics, Imperial College, London
(Ramsay Memorial Fellow, now at University College, Oxford)

Communicated by W. Jevons, 4 November 1939

ABSTRACT. In continuation of the work on ultra-violet band systems of the diatomic group-IV(b)+VI(b) molecules, the band spectrum of SnTe has been obtained in emission by a heavy-current uncondensed discharge through a mixture of Sn, Te and Al in a silica discharge tube. About 45 bands have been observed and measured in the region λλ3575 to 4235, of which 26 have been assigned to a main system A→X. Band heads of this system are represented by the equation $\nu_{head}=25451{\cdot}7+(177{\cdot}4u'-0{\cdot}13u'^2)-(263{\cdot}7u''-1{\cdot}1u''^2)$. Bands on the short-wave-length side of this system appear to belong to an incompletely developed system involving x, which is probably the ground state. Heads increasingly distant from their respective system-origins become increasingly diffuse owing to the vibrational isotope effect, which is briefly described. The relations of the electronic and vibrational constants of SnTe to those of the analogous oxides, sulphides, selenides and tellurides of Sn and Pb are reviewed, together with the problem of the correlation of the upper states of these molecules.

§1. INTRODUCTION

The present paper represents a further instalment of a programme of work undertaken to correlate spectroscopic data on the diatomic molecules MX, where M and X are a group-IV (b) and a group-VI (b) atom respectively. For the group-IV (b) elements close to Sn, data are available for all four Ge compounds, GeS (in absorption), GeO, GeSe and GeTe (in emission); and data on all four Pb compounds are available (in absorption, and PbO also in emission). Some particulars of six of these eight compounds are to be found in a previous paper of the present series[1], where bibliographic references to the several investigations are also given. The remaining two, GeSe and GeTe, are described in a forthcoming paper[2].

We examine first the results which have been obtained from an investigation of the closely related compounds PbSe, PbTe and SnSe. Systems due to these molecules have been observed in absorption by Walker, Straley and Smith[3].

(i) PbSe. In the region between λ4000 and λ6000, three band systems have been observed in the vapour of PbSe, analysed and designated as A←X, B←X and C←X systems. Bands evidently belonging to a D←X system and to an E←X system were also observed, but these were not sufficiently well developed to make

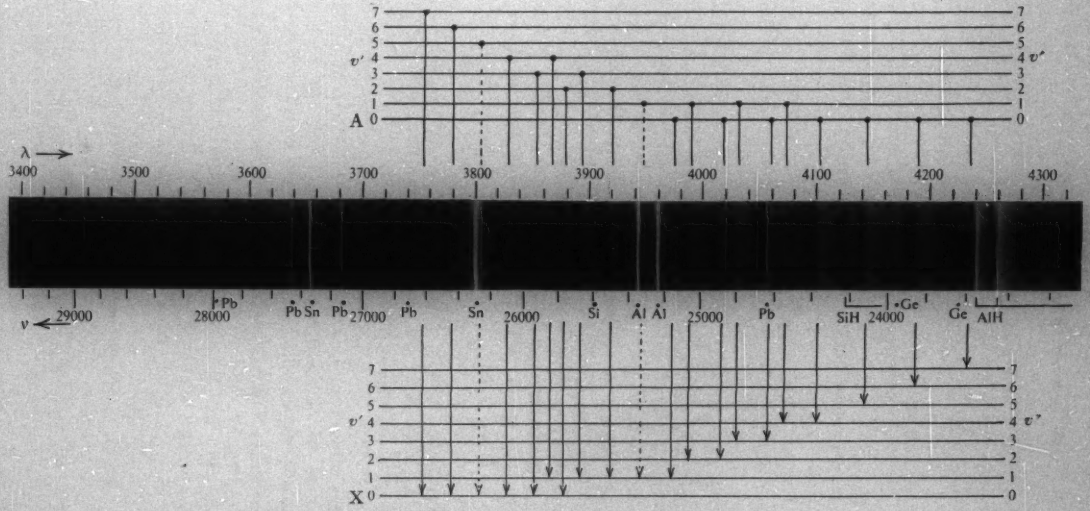


Figure 1. SnTe bands in emission (grating first order), and vibrational analysis of system A→X.

To face page 980

it possible to analyse them. The B←X system was the most completely developed of the systems.

(ii) PbTe. Three systems of bands were also observed in PbTe, but only two of them were analysed. These systems were regarded as corresponding to the A←X and to the B←X systems of PbSe, being displaced toward the red, when compared to the respective systems of PbSe.

(iii) SnSe. Three imperfectly developed systems of bands have been observed and analysed. They are similar to the three systems observed in PbSe, and are designated as the A←X, B←X and C←X systems.

CTe and SnTe are the only molecules of the whole group of twenty for which no data have yet been published. Attempts on the former have been entirely unsuccessful [4]. For the latter, one well-developed system of bands and a portion of a second system have been obtained in emission and are now to be described.

§ 2. ISOTOPE EFFECTS AND VIBRATIONAL ANALYSIS

The source used for the production of the band spectrum of SnTe was a heavy-current uncondensed discharge through a mixture of Al, Sn and Te contained in a silica tube of the type shown in figure 3 of a previous paper [1]. The electrical conditions and the spectrograph employed were the same as those used in other investigations in this series [1], [2], [4], [5].

The emission spectrum attributed to SnTe consists of red-degraded bands extending from λ3570 to λ4240, and is shown in the reproduction of a spectrogram in figure 1. It lies between the strong bands of the SiTe ultra-violet system [5] and the AlH bands at λ4241 and λ4259. The SiH band at λ4142 is also weakly developed. Atomic lines present are due to Pb, Sn and Si. In view of the results obtained by Walker, Straley and Smith, who record, for example, that SnSe presented considerable experimental difficulties by reason of the prominence of Se_2, SnO and SeO impurities, the cleanness of the present spectrograms taken in emission is a strong point in favour of this method.

Inspection of the plate shows that the heads of bands near the middle of the system are fairly sharp, whilst those of the outer bands become more and more diffuse as the distance from the centre increases. It is evident that this phenomenon is of isotopic origin, and it would perhaps be well to discuss briefly the isotope effects before presenting the band-head data and vibrational analysis.

The *International Table of Stable Isotopes* [6] shows 10 and 7 isotopes of Sn and Te respectively; accordingly there are 70 isotopic molecules of SnTe. This number is reduced to 24 if we ignore those whose abundance is 10 per cent or less in comparison with the most abundant molecule $^{120}Sn^{130}Te$. The relative intensities and positions of the heads in a given type (say R-branch heads) in a hypothetical fully-resolved band is shown in figure 2, in which the vertical lines are spaced on a horizontal scale of coefficients $(\rho-1)$ of vibrational displacements from the band-head due to the most prominent molecule, and their heights are proportional to relative abundance and therefore to intensity; the numerical data for the

construction of the diagram are readily calculated from the known mass-numbers and abundances of the two elements. Intensities of 10 per cent or less have been omitted for simplicity. The diagram contains one pair of superposed band-heads with identical displacement coefficients and there are also other pairs and groups of heads whose coefficients are so nearly equal that there is no possibility of their being resolved, their intensities being therefore to a large extent additive. In a given band (i.e. for given values of v' and v'') the band-head (*a*) due to the most abundant molecule has probably been identified and measured where it lies on the low-ν side of the system origin, for it is the strong head of highest ν; in a band on the high-ν side of the system origin, however, for which the diagram must be

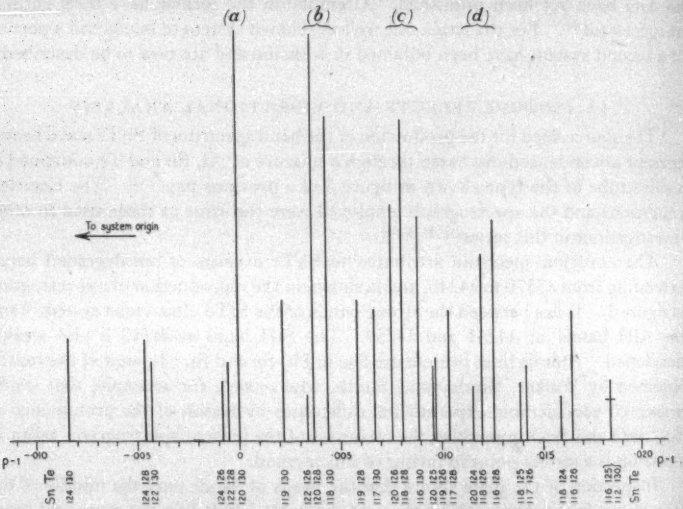


Figure 2. Theoretical isotope effect in an SnTe band.

laterally inverted as in a mirror-image, even the head (*a*) may not have been identified with certainty, in view of the fact that the bands degrade to low ν-values. The diagram indicates the possibility of measuring, in favourable cases, the maxima (*b*), (*c*) and (*d*) as diffuse blends, separated from (*a*) by roughly 0·004 ($\nu-\nu_e$), 0·008 ($\nu-\nu_e$) and 0·012 ($\nu-\nu_e$), as calculated for $^{120}Sn^{128}Te$, $^{118}Sn^{128}Te$ and $^{118}Sn^{126}Te$; but with the dispersion employed, owing to the smallness of the intervals between (*a*), (*b*), (*c*) and (*d*), and to the presence of intervening bands of considerable intensity, even this was not found possible, and the centre of gravity of the pattern only was measured.

For the derivation of the vibrational constants, therefore, the data for the sharper heads near the system origin have been given the greater weight, since it is only

in these bands that the measurements of blended heads are likely to apply even approximately to the most abundant species ^{120}Sn ^{130}Te. Deslandres arrangements of the data for the two systems are shown in tables 1 and 2.

Table 1. Deslandres arrangement of data for the A→X system of SnTe

Italics denote wave-lengths in air (I.A.);
Large Roman type, wave-numbers in vacuo ($cm.^{-1}$);
Small Roman type, wave-number differences (vibrational frequencies in upper and lower electronic states).

v' \ v''	0	1	2	3	4	5	6	7
8	26809·6 *3728·95*	26553·0 *3764·98*						
	172·6	177·4						
7	26637·0 *3753·12*	26375·6 *3790·31*						
	167·8							
6	26459·2 *3778·34*							
5	‡	26030·4 *3840·57*						
		176·1						
4	26116·3 *3827·94*	25854·3 *3866·74*						
	176·3	176·3						
3	25940·0 *3853·96*	25678·0 *3893·28*						
	178·3	176·8						
2	25761·7 *3880·63*	25501·2 *3920·27*				24469 *4085·6*	24216 *4128·3*	
		176·8						
1		25324·4 *3947·65*	25065·1 *3988·49*	24808·1 *4029·81*	24551·4 *4071·94*	†	*	23782 *4203·7*
		177·1	177·1	177·2	175·7			175
0		25147·3 *3975·45*	24888·0 *4016·87*	24630·9 *4058·80*	24375·7 *4101·29*	24117 *4145·3*	23862 *4189·6*	23607 *4234·8*

* Band probably present, but head not measurable on account of (0, 5) band.
† Band probably present, but head not measurable on account of (0, 4) band.
‡ Band probably present, but head masked by Sn line (λ3801·03).

(i) *System* A→X. The expression

$$\nu_{\text{head}} = 25451{\cdot}7 + (177{\cdot}4\,u' - 0{\cdot}13\,u'^2) - (263{\cdot}7\,u'' - 1{\cdot}1\,u''^2),$$

although leaving rather large residuals associated with the heads having $v''=6$ and 7, represents well the fifteen heads nearest to the system-origin.

(ii) *System* B→X. The arrangement of the bands as set out in table 2 is very tentative, but no better scheme has suggested itself. No attempt at mathematical representation of the data has been made in view of the uncertainty in this analysis.

Table 2. Deslandres arrangement of data for the B→X system of SnTe

Italics denote wave-lengths in air (I.A.);
Large Roman type, wave-numbers in vacuo (cm.$^{-1}$);
Small Roman type, wave-number differences (vibrational frequencies in upper and lower electronic states).

v'	v''→0	1	2	3	4	5
3	27949·5 *3576·86*		27427·8 *3644·90*			
	119·2		116·0			
2	27830·3 *3592·19*	260·8 27569·5 *3626·16*	257·7 27311·8 *3660·38*			
		122·0	118·6			
1		27447·5 *3642·29*	254·3 27193·2 *3676·34*	255·4 26937·8 *3711·20*	255 26683 *3746·7*	254 26429 *3782·6*
			129·4	128·2	130	
0			27063·8 *3693·92*	254·2 26809·6* *3728·95*	256·6 26553·0† *3764·98*	

* Also the (8, 0) band of A→X system.
† Also the (8, 1) band of A→X system.

It does appear, however, if the present scheme is partially correct, that the same lower state is involved in both these systems, and it is probable that this is the ground state of SnTe.

(iii) In addition to the bands listed above, four weaker bands have been measured at λλ4037·7, 4013·7, 3997·3 and 3972·6. These may form a fragment of a third system, thus:

25165	257	24908
155		148
25010	250	24760.

The energy of dissociation in the lower state of the A→X system is found by linear extrapolation using the expression

$$D_0=(\omega_e-x_e\omega_e)/(4x_e\omega_e\times 8106)$$

to be 1·9 ev. This bears a satisfactory relation to the values of the energy of dissociation of SnO, SnS and SnSe in their ground states, viz. 5·7, 5·0 and 2·7 ev. It seems probable that there is a fairly regular decrease in this constant as the molecular weight increases. A more detailed discussion of the variation of this constant is deferred until a later paper[2].

§3. DISCUSSION

The general behaviour of the vibrational constants of the band systems of the compounds of C, Si and Ge with the group-VI (b) elements when subjected to the empirical tests discussed in a previous paper[1] leaves little doubt that these systems are, in an arithmetical sense, analogous, although they may be of differing theoretical type (e.g. $^1\Pi \rightarrow {}^1\Sigma$ or $^1\Sigma \rightarrow {}^1\Sigma$). In the case of the band systems of Sn and Pb, however, the problem is by no means so simple, since for these molecules at least three systems are known (although analyses may be uncertain, owing to incompleteness of the experimental data), while for PbO and PbS, five systems have been postulated. Moreover, these systems all lie in the same region of the spectrum, stretching from the blue to the near ultra-violet, and different systems of the same molecule are frequently superimposed. With the exception of their degree and intensity of development, which varies somewhat with experimental conditions, they are all, superficially at least, similar in character.

A correlation of the systems of SnO, SnS, PbO and PbS was suggested by Howell[7], based on the values taken by the ratios ω_e'/ω_e'' and $\omega_e(\text{sulphide})/\omega_e(\text{oxide})$, but the extension of this correlation to include the systems of PbSe, PbTe and SnSe gave inconclusive results[1]. It seemed therefore that although Howell's original correlation was plausible, some caution was necessary in attaching a more general significance to the functions from which the correlation was derived.

It is well known that the physical constants of iso-electronic molecules take values which often approach one another very closely: among the band systems under consideration, one, namely the B←X system of SnS, may be provisionally identified as being nearly related to the band systems of the C, Si and Ge compounds. SnS, with 66 electrons, is iso-electronic to SiTe, and the values of ν_e and ω_e' for their ultra-violet systems are 28338 and 28664, and 332 and 336 respectively.

In the paper on SiS and its related molecules[1] the variation of the electronic energy E_e with the ionization potentials, I_M and I_X, of the component atoms was studied by observing the values taken by the function $I_M I_X/E_e$. The conclusion that "in this group the molecular electronic energies E_e have the same general trends as the ionization potentials though they are not strictly proportional thereto" was necessarily rather indefinite, but since the appearance of that paper, other data have been obtained (on band systems of CSe[4], SiSe and SiTe[5], GeSe and GeTe[2], and SnTe). Table 3 gives the values of this ratio for the C, Si and

Table 3. Values of $I_M I_X/E_e$ for the group-VI (b) compounds of C, Si and Ge

	$_8$O	$_{16}$S	$_{34}$Se	$_{52}$Te
$_6$C	18·9*	24·0*	24·5	—
$_{14}$Si	20·8*	19·4	19·3	20·6
$_{32}$Ge	23·5†	20·5	20·5	21·0

* Upper state $^1\Pi$.

† Upper state $^1\Sigma$ (Sen Gupta[8]).

Ge compounds. These are seen to be sensibly constant throughout, with the exception of the rather high values for CS, CSe and GeO (upper state $^1\Sigma$). The value taken by the B←X system of SnS, identified as above, is 21·5. On this evidence it seems reasonable to use the relative constancy of this function as a test for correlation: accordingly the values of $I_M I_X/E_e$ for the compounds of Sn and Pb are set out in table 4. From the results of this test the excited states of the latter compounds to be correlated with those of the C, Si and Ge compounds are evident, if we accept the hypothesis that the electronic energies of corresponding states of related molecules are proportional to the products of the ionization potentials of the component atoms. They are the states D of SnO, B of SnS, C of SnSe, A of SnTe, F of PbO and E of PbS.

Table 4. Values of $I_M I_X/E_e$ for the group-VI (b) compounds of Sn and Pb. (The figures in parentheses underneath the chemical symbols denote the electron numbers. The letters A, B, C . . . are the accepted designations of the relevant excited states.)

SnO (58)	SnS (66)	SnSe (84)	SnTe (102)	PbO (90)	PbS (98)	PbSe (116)	PbTe (134)
	18·4 C						
21·8 D	21·5 B	20·9 C	20·8 A	22·7 F	20·8 E		
		24·9 B				24·4 C	
27·1 A†	26·3 A			26·8 L	28·2 B	27·1 B	27·2 B
31·8 B		29·1 A		32·6 C	32·7 A† [9]	30·4 A	32·8 A
				35·4 B†			

† Upper state $^1\Sigma$.

From this table have been omitted the A state of PbO, lying below the B state, and the C and D states of PbS, since for the systems C←X and D←X, bands with v'' values 0 and 1 only are known, and the positions of their system origins is very uncertain.

It must be noted that this new correlation of excited states destroys the smoothness of the oxide and sulphide curves of $1/\omega_e'$ against n previously found for the states chosen in the former study [1], although the Sn and Pb curves for this ratio are still without inflexion. Thus there appears to be some contradiction between the provisional correlations of states obtained from the $1/\omega_e'$ curves, and from the ratio $I_M I_X/E_e$. It is hoped to extend the examination of these functions to other homologous groups of molecules, when it may be possible to pronounce more definitely on the relative values, if any, of these two methods of classification.

It may possibly be objected that the process of seeking an empirical relation as a basis for the classification of the excited states of these molecules is of little value as no account of their theoretical relations is given. Unfortunately, very little is known as to the theoretical types of these states (in the case of the heavier molecules rotational analysis will be exceedingly difficult), and it seems that a tentative scheme of correlation, such as that above, is not without interest. The problem will not,

of course, be settled until these types are satisfactorily determined, and until a theory of the relative probabilities of the electronic transitions in these systems has been evolved. The immediate requirement, it is suggested, is the accumulation of more experimental data from spectrographic studies in emission (where these have not already been carried out) and from rotational analysis, wherever this is possible.

§ 4. ACKNOWLEDGEMENT

The author takes pleasure in recording once more his indebtedness to Dr. W. Jevons for helpful discussion throughout the course of this work, and for assistance in the preparation of this paper.

REFERENCES

(1) Barrow, R. F. and Jevons, W. *Proc. Roy. Soc.* A, **169**, 49 (1938).
(2) Barrow, R. F. and Jevons, W. Forthcoming paper on GeSe and GeTe.
(3) Walker, J. W., Straley, J. W. and Smith, A. W. *Phys. Rev.* **53**, 140 (1938).
(4) Barrow, R. F. *Proc. Phys. Soc.* **51**, 989 (1939).
(5) Barrow, R. F. *Proc. Phys. Soc.* **51**, 267 (1939).
(6) " International Table of Stable Isotopes," *Rev. Sci. Instrum.* **7**, 334 (1936).
(7) Howell, H. G. *Proc. Roy. Soc.* A, **153**, 683 (1936).
(8) Sen Gupta, A. K. *Proc. Phys. Soc.* **51**, 62 (1939).
(9) Bell, H. and Harvey, A. *Proc. Phys. Soc.* **50**, 427 (1938).

NOTE ON THE SECOND POSITIVE BAND-SYSTEM OF NITROGEN

By R. C. PANKHURST, A.R.C.S., B.Sc.,
Imperial College, London, S.W.7*

Received 14 *November* 1939

THE second positive band-system of nitrogen is of frequent occurrence as an impurity in the spectra of discharge-tube sources. Although an extensive study has been made of the rotational structure of individual bands, there appears to be no record of the wave-lengths of the heads of the whole system which can be relied upon to 0·1 A. New measurements have therefore been made, medium and large quartz spectrographs being used for photographing the ultra-violet region, and the first and second orders of a 6-metre concave grating (11,000 lines/inch) for the visible region.

λ	*I*	(*v'*, *v''*)	λ	*I*	(*v'*, *v''*)
4976·5	00	(4, 11)	3805·0	6	(0, 2)
			3755·5	6	(1, 3)
4916·8	0	(1, 7)	3710·6	5	(2, 4)
4814·8	1	(2, 8)	3671·9	4	(3, 5)
4723·5	1	(3, 9)	3641·7	3	(4, 6)
4649·(6)	1	(4, 10)			
			3577·0	7	(0, 1)
4667·4	0	(0, 5)	3536·7	6	(1, 2)
4574·2	2	(1, 6)	3500·4	4	(2, 3)
4490·(2)	3	(2, 7)	3469	[Very weak]	(3, 4)
4416·4	3	(3, 8)	3446	[Very weak]	(4, 5)
4355·0	2	(4, 9)			
			3371·4	8	(0, 0)
4343·(7)	3	(0, 4)	3338·(5)	4	(1, 1)
4269·7	4	(1, 5)	3309·(1)	4	(2, 2)
4200·(5)	4	(2, 6)	3285·1	5	(3, 3)
4141·8	4	(3, 7)	3268·0	5	(4, 4)
4094·9	3	(4, 8)			
			3159·1	7	(1, 0)
4059·4	6	(0, 3)	3135·8	6	(2, 1)
3998·4	5	(1, 4)	3116·6	5	(3, 2)
3943·1	5	(2, 5)	3104·0	4	(4, 3)
3894·8	4	(3, 6)			
3857·8	3	(4, 7)	2976·8	5	(2, 0)
			2962·0	5	(3, 1)
			2953·2	5	(4, 2)
			2819·8	1	(3, 0)
			2814·3	1	(4, 1)

* Now a member of the staff of the National Physical Laboratory, Teddington.

The source was an uncondensed 2000-volt A.C. discharge through a slow stream of nitrogen. Several plates were also taken with air instead of nitrogen; with this source, in addition to the development of the γ bands of NO, it was observed that the weak (1, 1) and (2, 2) bands of the second positive system of N_2 were enhanced relatively to the (0, 0) band, and the red and yellow bands of the first positive system weakened with respect to the violet bands of the second positive. The discharge through nitrogen was consequently bronze in colour, whereas with air it was bluish-pink.

The second positive system arises from a $^3\Pi \rightarrow {}^3\Pi$ transition, forming triple-headed bands degraded to the violet. It is the head with longest wave-length that has been measured in each case. The order of accuracy to which the results are given (0·1 A.) should be sufficient for identification of the system. The intensities are eye-estimates from photographs of discharges through air. In the case of those bands whose rotational structure has been analysed[1], the heads as extracted from the analyses agree to within 0·1 A. with the present results. Measurements made on parts of the system by Hermesdorf[2] and by Lawton[3], when reduced to international angstroms, also agree to the same order of accuracy.

REFERENCES

(1) Zeit, P. *Z. wiss. Photogr.* **21**, 1 (1921);
Hulthèn, E. and Johansson, G. *Ark. Mat. Astr. Fys.* **18**, 28 (1924);
Hulthèn, E. and Johansson, G. *Z. Phys.* **26**, 308 (1924);
Lindau, P. *Z. Phys.* **26**, 343 (1924); *Z. Phys.* **30**, 187 (1924);
Coster, D., Brons, F. and van der Ziel, A. *Z. Phys.* **84**, 304 (1933);
Guntsch, A. *Z. Phys.* **86**, 262 (1933).
(2) Hermesdorf, P. Dissert., Bonn (1902); *Ann. Phys., Lpz.*, **11**, 161 (1903).
(3) Lawton, E. E. *Amer. J. Sci.* **24**, 101 (1907).

THE COLOUR OF PIGMENT MIXTURES

By D. R. DUNCAN, B.Sc., Ph.D.,
Private Laboratory, 50, Buckleigh Avenue, London, S.W.20

Communicated by Dr. R. F. Hanstock

ABSTRACT. It is shown that the colours produced by mixing pigments in any given medium may be deduced from formulae involving two constants for each pigment. These constants, which vary with the wave-length of the incident light, measure respectively the absorbing power of the pigment for light and its scattering power. Relative values of these constants for a series of pigments may be deduced from their reflectivities at different wave-lengths and the reflectivities of one mixture of each pigment with one standard pigment. The relative values so obtained enable the reflexion curves, and hence, if desired, the colours, of any mixture of two or more of the pigments investigated to be predicted correctly.

§ 1. INTRODUCTION

It is a well-known fact that the colours produced by mixing coloured pigments are different from those produced by mixing lights of the same colours. Whereas in the latter case the colours produced follow known laws, no method has hitherto been known for predicting quantitatively the colours produced by mixing pigments. In the present paper, however, it will be shown that the colours of such mixtures are dependent in a relatively simple way on the optical constants of the pigments and may be accurately predicted if these are known.

The nature of the colours produced by the mixture of coloured lights was examined by a series of workers and led to the modern methods of colorimetry and systems of colour specification typified by the C.I.E. system[1]. It is now possible to predict with certainty the colour which will be produced on mixing any number of lights of known colour in any given proportions. Hence, it is possible to deduce the colour of any light, given its spectral composition. The problem of the prediction of the colours of pigment mixtures therefore reduces to the purely physical one of determining the relationship between the reflexion spectrum of the pigment mixture and the composition of the mixture. Since the different wave-lengths may be considered independently and the effect of each wave-length on the colour finally integrated over the range of the visible spectrum, it is sufficient to discover the general law relating the reflectivity of a pigment mixture for any one wave-length to the optical constants of the component pigments and the proportions in which they are mixed. In the present paper, such a relationship will be evolved and shown to apply to a series of pigment mixtures.

§ 2. THEORETICAL

It is commonly stated that when coloured lights are mixed the effect is one of additive mixture, but that when coloured pigments are mixed the effect is subtractive, i.e., that the absorptions produced by the separate pigments are added together to produce the colour of the mixed pigment. Thus, the green colour observed in many cases when a blue pigment and a yellow pigment are mixed is attributed to the fact that green is the only colour reflected by both pigments; light of any other colour is absorbed either by the blue pigment or by the yellow pigment, and is therefore absorbed by the mixture. This generalization, however, is only an approximation to the truth, and lacks the quantitative precision necessary for the prediction of the exact shade of colour which will be produced when the pigments are mixed in varying proportions.

In the case where two substances in solution are mixed (in the absence of chemical reaction), the colour of the mixture may be deduced from the extended form of Beer's law

$$I/I_0 = e^{-(\epsilon_A c_A + \epsilon_B c_B + \ldots)h} \quad \ldots\ldots(1),$$

where I is the intensity of the transmitted light, I_0 that of the incident light, h is the thickness of solution traversed by the light, and $\epsilon_A, \epsilon_B, \ldots$ are the extinction coefficients and $c_A, c_B, \ldots$ the concentrations of the substances A, B, etc. A similar relationship holds when light is passed through a series of coloured, transparent filters. When, however, the coloured substances are mixed as dry powders or as powders in suspension in a medium, as in the mixing of artists' colours, the manufacture of paints, and the colouring of cement, rubber, paper, etc. by means of pigments, this law no longer applies.

Various formulae have been evolved for the reflecting (more accurately, remitting) and transmitting powers of suspensions of solid particles, and those recently put forward by Amy, Sannié and Sarraf [2] are now found to form a basis from which, making certain assumptions, it is possible to deduce an equation by means of which the colours of paints containing a mixture of different pigments may be predicted correctly.

According to these investigators, for a layer of translucent material,

$$\frac{I}{I_0} = \frac{2\sqrt{\alpha(\alpha+2\delta)}e^{-h\sqrt{\alpha(\alpha+2\delta)}}}{\alpha+\delta+\sqrt{\alpha(\alpha+2\delta)}-[\alpha+\delta-\sqrt{\alpha(\alpha+2\delta)}]e^{-2h\sqrt{\alpha(\alpha+2\delta)}}} \quad \ldots\ldots(2)$$

and

$$R = \frac{\delta[1-e^{2h\sqrt{\alpha(\alpha+2\delta)}}]}{\alpha+\delta+\sqrt{\alpha(\alpha+2\delta)}-[\alpha+\delta-\sqrt{\alpha(\alpha+2\delta)}]e^{-2h\sqrt{\alpha(\alpha+2\delta)}}} \quad \ldots\ldots(3),$$

where R is the ratio of the intensity of remitted (so-called reflected) light to that of the incident light, h is the thickness of the layer of material, and α and δ are constants for the material for any given wave-length, but, in general, vary with the wave-length of the incident light, α being the extinction coefficient as defined by the relationship: light absorbed by an element of the film of thickness $dh = I_0\alpha\,dh$, while δ is the coefficient of scatter, similarly defined by the relationship: light remitted by the element due to scattering $= I_0\delta\,dh$. Similar expressions have been proposed by Neugebauer [3].

When the thickness is such that any further increase therein produces no further change in colour by reflected light (i.e., in R), a condition which is fulfilled when $e^{-2h\sqrt{a(a+2\delta)}}$ becomes negligible compared to 1, these expressions reduce to

$$\frac{I}{I_0}=\frac{2\sqrt{a(a+2\delta)}e^{-h\sqrt{a(a+2\delta)}}}{a+\delta+\sqrt{a(a+2\delta)}} \qquad \ldots\ldots(4)$$

and

$$R=\frac{\delta}{a+\delta+\sqrt{a(a+2\delta)}} \qquad \ldots\ldots(5).$$

In the present paper attention will be confined mainly to the application of equation (5) to paint films. Amy and collaborators confirmed the truth of various simplified expressions deduced from (4) and (5) for cases in which δ^2 is negligible compared with a^2. In particular, they studied the change in R on adding (to paper or white paint) a small proportion of an intensely coloured substance, which was assumed to cause an increase in a and to have no appreciable effect on the value of δ. In the present paper, the general case is considered, in which a and δ may have any values and the proportion and nature of the pigments added may be such as to cause considerable variation in both a and δ.

Attention is confined initially to the case where the medium is clear and colourless.

For the present purpose, equation (5) may be more conveniently written in the form

$$\delta/a=2R/(1-R)^2 \qquad \ldots\ldots(6),$$

which is obtained by solving (5) as a quadratic in δ, the other root being $\delta=0$.

δ can have the value zero only in certain special cases, viz., when the film is unpigmented, when the pigment dissolves completely in the medium (in which case it is more suitably described as a dye or stain), and when the refractive index of the pigment is identical with that of the medium. In these cases, δ being zero, equation (4) reduces to $I/I_0=e^{-hx}$ (Beer's law) and (5) becomes $R=0$.

Since the expression $2R/(1-R)^2$ will occur frequently throughout the work, it is convenient to refer to it as the " reflectivity function " of the paint and to denote it by the letter Φ:

$$\Phi=2R/(1-R)^2 \qquad \ldots\ldots(7).$$

If it is assumed that for a paint containing a mixture of pigments, viz., c_A parts of pigment A, c_B parts of pigment B, etc., the values of a and δ each follow a simple additive law, i.e.,

$$a_M=c_Aa_A+c_Ba_B+c_Ca_C+\ldots$$

and

$$\delta_M=c_A\delta_A+c_B\delta_B+c_C\delta_C+\ldots$$

where a_M and δ_M are the values of a and δ for the mixed paint, a_A and δ_A are the values for a paint containing only A, and so on, then

$$\Phi_M=\frac{c_A\delta_A+c_B\delta_B+c_C\delta_C+\ldots}{c_Aa_A+c_Ba_B+c_Ca_C+\ldots} \qquad \ldots\ldots(8).$$

In the practical application of this expression, c_A, c_B, etc. may conveniently be taken as the percentages of each pigment in the pigment mixture, ignoring the medium, since variation of the proportion of medium, except when excessive, has no influence on the value of Φ.

In the special case of the binary mixture of a white pigment (or other white substance, such as paper pulp, cement, etc.) A, with a coloured pigment, B, the value of α_A may be taken as zero to a close approximation, when (8) simplifies to

$$\Phi_M = (c_A\delta_A + c_B\delta_B)/c_B\alpha_B,$$

whence
$$\Phi_M = \frac{c_A}{c_B}\cdot\frac{\delta_A}{\alpha_B} + \frac{\delta_B}{\alpha_B} \qquad \ldots\ldots (9),$$

or
$$\Phi_M = \frac{c_A}{c_B}\cdot\frac{\delta_A}{\alpha_B} + \Phi_B \qquad \ldots\ldots (10).$$

Hence the reflectivity function for a series of paints containing different ratios of A to B is a linear function of that ratio, a relationship which has been confirmed experimentally for numerous mixtures. The gradient of this line gives the value of the ratio δ_A/α_B and the reflectivity function of the paint containing the undiluted pigment B gives the value of the ratio δ_B/α_B. Hence, α_B and δ_B may be evaluated in terms of δ_A. Similarly, the relative values of α and δ for a series of pigments may be deduced from the reflectivities of each pigment and of one mixture of each pigment with the standard white pigment. Instead of mixing the pigment with the standard white pigment, it may be mixed with any pigment which has already been compared with the white pigment in this way, equation (8) being then used for deducing the values of α and δ for the pigment under test. In order to obtain reasonable accuracy, the two pigments must differ considerably in reflectivity.

From the relative values of α and δ so obtained, it is possible to calculate the reflectivity function of a paint containing a mixture of any number of these pigments in any proportions, by means of equation (8). The reflectivity follows from equation (7), and having calculated the reflectivity for different wavelengths, the colour of the paint may be calculated therefrom[1].

The values for the reflectivities of paints containing different pigment mixtures deduced in this way agree within the limits of experimental error with those found by direct measurement.

§ 3. EXPERIMENTAL METHOD

The paints for examination were made up by grinding the required weights of each pigment together with a concentrated aqueous solution of gum arabic by hand in a mortar until complete homogeneity, as indicated by the absence of change in colour on further grinding, was obtained. The resultant paste was spread thickly and evenly on cardboard and allowed to dry. A disc was then cut out of this coated cardboard and mounted concentrically on top of a pair of larger

discs, one black and one white, which were interlocked so that the proportions of black and white exposed could be varied at will (Maxwell's discs). The discs were illuminated by light incident at 45° and were viewed normally through a system of filters transmitting a narrow band of the spectrum. The angle of the black and white sectors was varied until a match was obtained on revolving the discs, the mean of a number of readings being taken.

The black and white discs used were coated respectively with lampblack and with white lead in gum arabic. The relative reflectances of these discs were determined photometrically, i.e., by measuring the relative intensities of illumination on the two discs necessary to produce a match when viewed through the different combinations of filters.

The results were calculated relative to the reflectivity of the white lead as standard, this substance having, in the gum-arabic medium, the highest reflectivity at all wave-lengths of all the pigments examined. The same sample of white lead was employed throughout the work and was shown by analysis to consist of pure basic lead carbonate of the formula $Pb(OH)_2, 2PbCO_3$. This pigment was used as the standard white pigment A referred to in the theoretical section.

It was found that a series of filters transmitting suitable bands could be obtained by combining Ilford Micro filters in pairs. The transmittances were determined by means of a Nutting photometer and the mean wave-lengths for the transmission bands calculated according to the formula:

$$\lambda_{\text{mean}} = \Sigma\left(\frac{I}{I_0} L\lambda\right) \Big/ \Sigma\left(\frac{I}{I_0} L\right),$$

where L is the visual luminosity at wave-length λ of the light used. The values of λ_{mean} are accurate only for the particular source of light used. This source was too poor in violet for accurate measurements to be taken in this region. The combinations of filters used are shown in table 1.

Table 1

Series of filter combinations used for isolating different bands of the spectrum

Ilford micro filter nos.	Colour	Mean wave-length (mμ.)
1+6	Blue	460
1+9	Blue	483
2+4	Green	530
3+5	Yellow	578
5+7	Orange-red	616
5+6	Red	675

§ 4. EXPERIMENTAL RESULTS

In illustration of the results obtained, the data for mixtures of zinc oxide and lampblack in yellow light will be considered in detail, and the results with other mixtures will then be reported rather more briefly.

For mixtures of this type (one white and one coloured pigment), equation (10) would be expected to apply. This may be written

$$\frac{c_A/c_B}{\Phi_M-\Phi_B}=k,$$

where k is a constant.

The data obtained are given in table 2. It will be seen that there is a close agreement between the observed and calculated values of R. Since a small error in R introduces a relatively large error in Φ and hence in k (e.g., an error of 0·01 in R corresponds to a 6 per cent error in Φ and k when $R=0{\cdot}50$), it will be seen that the value of k is constant within the limits of accuracy attainable experimentally. If c_A/c_B is plotted against $\Phi_M-\Phi_B$, it will be found that a close approximation to a straight line is obtained.

Table 2. Reflectivity data for mixtures of zinc oxide and lampblack in a gum-arabic medium for yellow light of mean wave-length 578 mμ.

c_A	c_B	c_A/c_B	$R_{obs.}$	Φ_M	$\Phi_M-\Phi_B$	k	$R_{calc.}$
100	0	∞	1·00	∞	∞	—	1·00
98·8	1·2	82·3	0·60	7·50	7·47	11·0	0·60
97·0	3·0	32·3	0·43	2·65	2·62	12·3	0·45
95·0	5·0	19·0	0·36	1·76	1·73	11·0	0·36
93·1	6·9	13·5	0·31	1·30	1·27	10·6	0·31
90·0	10·0	9·0	0·25	0·89	0·86	10·5	0·24
87·0	13·0	6·7	0·21	0·674	0·642	10·4	0·20
80·0	20·0	4·00	0·14	0·379	0·347	11·5	0·14
73·0	27·0	2·70	0·11	0·278	0·246	11·0	0·11
60·0	40·0	1·50	0·075	0·174	0·142	10·6	0·072
0	100	0	0·016	0·032	0·000	—	0·016
Mean value of k (used in calculating R)						11·0	

According to the theory, $k=\alpha_B/\delta_A$ and $\Phi_B=\delta_B/\alpha_B$. Hence, α for lampblack is $11{\cdot}0\delta_{ZnO}$ and δ is $0{\cdot}35\delta_{ZnO}$ for yellow light ($\lambda=578$ mμ.). Corresponding values for lampblack relative to $\delta_{\text{white lead}}$ are similarly found to be $\alpha=4{\cdot}3\delta_{\text{white lead}}$ and $\delta=0{\cdot}14\delta_{\text{white lead}}$. The observed values of R for mixtures of lampblack and white lead, and the values calculated assuming the above figures, are given in table 3.

Table 3. Reflectivity data for mixtures of lampblack and white lead in yellow light ($\lambda=578$ mμ.)

Per cent lampblack	$R_{observed}$	$R_{calculated}$
0	1·00	1·00
5	0·49	0·48
10	0·34	0·35
20	0·24	0·23
30	0·18	0·16
50	0·10	0·10
100	0·015	0·015

Hence we have $\delta_{\text{lampblack}}=0{\cdot}35\delta_{\text{ZnO}}=0{\cdot}14\delta_{\text{white lead}}$, whence $\delta_{\text{ZnO}}=0{\cdot}40\delta_{\text{white lead}}$. The observed ratio of the reflectivities of the zinc oxide and white lead used is as 0·96 to 1·00 at $\lambda=578$ mμ. Hence, if the white lead is taken as the standard of whiteness ($R=1$, whence $\alpha=0$), then Φ_{ZnO} is 1200. Hence, $\delta_{\text{ZnO}}/\alpha_{\text{ZnO}}=1200$, and therefore

$$\alpha_{\text{ZnO}}=\delta_{\text{ZnO}}/1200=0{\cdot}40\delta_{\text{white lead}}/1200=3\times10^{-4}\,.\,\delta_{\text{white lead}}.$$

Whilst absolute values of α and δ could be obtained from opacity measurements of thin films, these relative figures are all that is required for the calculation of the colours of pigment mixtures by reflected light. Relative values on this basis ($\alpha_{\text{white lead}}=0$, $\delta_{\text{white lead}}=1$) will be written $[\alpha]$ and $[\delta]$ for purposes of distinction from the absolute figures. Since $\alpha_{\text{white lead}}$ is very small in comparison with $\delta_{\text{white lead}}$, $[\alpha]=\alpha/\delta_{\text{white lead}}$ and $[\delta]=\delta/\delta_{\text{white lead}}$ to a close approximation.

The values of $[\alpha]$ and $[\delta]$ obtained in this way for several pigments at a number of wave-lengths are given in table 4. It should be noted that the value of $\delta_{\text{white lead}}$ probably varies somewhat with the wave-length, so that the unit of measurement is not exactly the same at all wave-lengths, but this does not affect the accuracy with which the colours of mixtures may be predicted. It should also be stressed that these figures refer only to the particular samples of pigment used, in the gum-arabic medium employed. In particular, $[\delta]$ would be expected to vary with the particle size of the pigment and with the refractive index of the medium.

Table 4

Values of $[\alpha]$ and $[\delta]$ for a series of pigments.
Standard: white lead, $[\alpha]=0$, $[\delta]=1$, for all wave-lengths

		Blue 460 mμ.	Blue 483 mμ.	Green 530 mμ.	Yellow 578 mμ.	Orange-red 616 mμ.	Red 675 mμ.
Red lead	$[\alpha]$	0·54	0·64	0·61	0·08	3×10^{-3}	3×10^{-3}
	$[\delta]$	0·034	0·070	0·115	0·32	0·44	0·56
Lampblack	$[\alpha]$	4·5	4·5	4·7	4·3	5·0	5·9
	$[\delta]$	0·17	0·17	0·16	0·14	0·14	0·14
Zinc oxide	$[\alpha]$	4×10^{-3}	2×10^{-3}	6×10^{-4}	3×10^{-4}	2×10^{-4}	2×10^{-4}
	$[\delta]$	0·42	0·40	0·39	0·40	0·43	0·44
Lead chromate	$[\alpha]$	5·3	4·3	0·76	0·022	0·0092	0·0064
	$[\delta]$	1·00	0·93	1·27	2·07	2·06	1·84
Chromium oxide	$[\alpha]$	10·6	9·8	2·82	5·55	7·4	4·3
	$[\delta]$	3·3	2·9	3·2	2·11	2·5	2·6

In illustration of the accuracy with which the reflectivities of pigment mixtures may be predicted by means of these data, all the observed values of R for a series of binary mixtures of chromium oxide and lead chromate with one another and with white lead, two ternary mixtures of these three pigments, and one quaternary mixture with lampblack as the fourth pigment are compared in table 5 with the

figures calculated by means of equations (7), (8) and (10) using the values of [α] and [δ] given in table 4.

Table 5. Observed and calculated values of R for various pigment mixtures. The calculated values are given in parentheses

Pigments	% second pigment*	Blue 460 mμ.	Blue 483 mμ.	Green 530 mμ.	Yellow 578 mμ.	Orange-red 616 mμ.	Red 675 mμ.
Lead chromate + chromium oxide	0	0·10 (0·10)	0·09 (0·09)	0·35 (0·35)	0·87 (0·87)	0·91 (0·91)	0·92 (0·92)
	5	0·10 (0·10)	0·09 (0·09)	0·34 (0·34)	0·58 (0·59)	0·56 (0·55)	0·63 (0·62)
	20	0·10 (0·10)	0·10 (0·10)	0·30 (0·32)	0·37 (0·37)	0·32 (0·33)	0·40 (0·40)
	50	0·10 (0·11)	0·11 (0·11)	0·29 (0·30)	0·22 (0·22)	0·20 (0·20)	0·27 (0·27)
	80	0·12 (0·12)	0·12 (0·11)	0·29 (0·29)	0·19 (0·17)	0·16 (0·16)	0·23 (0·22)
	100	0·12 (0·12)	0·12 (0·12)	0·29 (0·29)	0·14 (0·14)	0·13 (0·13)	0·19 (0·19)
White lead + lead chromate	5	0·52 (0·50)	0·52 (0·52)	0·76 (0·76)	0·97 (0·96)	0·97 (0·97)	0·97 (0·97)
	10	0·37 (0·37)	0·40 (0·40)	0·69 (0·68)	0·95 (0·95)	0·96 (0·96)	0·97 (0·97)
	20	0·26 (0·27)	0·26 (0·28)	0·57 (0·58)	0·92 (0·93)	0·95 (0·94)	0·96 (0·96)
	30	0·19 (0·20)	0·19 (0·21)	0·52 (0·53)	0·89 (0·90)	0·92 (0·94)	0·94 (0·94)
	40	0·16 (0·17)	0·17 (0·17)	0·48 (0·48)	0·88 (0·89)	0·92 (0·93)	0·93 (0·94)
	50	0·16 (0·16)	0·15 (0·14)	0·45 (0·45)	0·88 (0·89)	0·91 (0·92)	0·92 (0·93)
White lead + chromium oxide	10	0·32 (0·29)	0·32 (0·32)	0·51 (0·50)	0·38 (0·38)	0·36 (0·34)	0·44 (0·43)
	20	0·21 (0·21)	0·23 (0·23)	0·39 (0·40)	0·27 (0·27)	0·25 (0·25)	0·33 (0·34)
	30	0·19 (0·18)	0·18 (0·18)	0·34 (0·37)	0·23 (0·23)	0·22 (0·21)	0·29 (0·29)
	50	0·15 (0·14)	0·14 (0·14)	0·30 (0·33)	0·18 (0·18)	0·16 (0·16)	0·24 (0·24)
White lead Lead chromate Chromium oxide	12·5 75·0 12·5	0·11 (0·10)	0·11 (0·13)	0·34 (0·38)	0·40 (0·44)	0·37 (0·39)	0·46 (0·47)
White lead Lead chromate Chromium oxide	80·0 10·0 10·0	0·24 (0·23)	0·26 (0·26)	0·47 (0·47)	0·43 (0·40)	0·39 (0·36)	0·47 (0·45)
White lead Lead chromate Chromium oxide Lampblack	50·0 25·0 12·5 12·5	0·14 (0·14)	0·14 (0·14)	0·26 (0·28)	0·27 (0·28)	0·23 (0·25)	0·26 (0·27)

* Percentage composition in the last three entries.

In calculating the theoretical value of R, it is necessary to solve the quadratic equation (7); of the two solutions obtained, the value to be accepted is the one lying between 0 and 1, the other root, which is greater than unity, representing a purely imaginary case, since the reflectivity cannot lie outside this range. In order to avoid the solution of numerous quadratic equations, the corresponding values of Φ and R are tabulated in table 6.

§ 5. DISCUSSION OF RESULTS

It will be seen that the experimental results confirm the theoretical reasoning advanced in § 2 and indicate that the colour-mixing properties of a pigment may be expressed in terms of two constants which are respectively measures of the light-absorbing and light-scattering properties of the pigment. These constants vary with the wave-length of the incident light, but are otherwise invariant in any one medium.

Whilst no attempt has been made in this paper to study quantitatively the effect of variation of the medium, it is to be expected that, whilst α is probably independent of the medium, δ will vary considerably with the refractive index of the medium and may also be affected by its pigment-dispersing properties, since if the pigment is present as large flocculates its scattering power will be less than if it is well dispersed. The use of a medium of high refractive index may also involve the introduction of a correction for reflexion at the air-medium interface. This reflexion is of two types: (1) external reflexion (specular reflexion), which is constant for all paints made up in the same medium, provided the proportion of pigment is not too high; and (2) internal reflexion, which would have the effect of increasing the apparent thickness but would probably have no effect on the colour at infinite thickness.

In the event of the medium being coloured, allowance could be made for this by considering it as an additional pigment with $\delta=0$ and α having a value depending on its colour (equal to its extinction coefficient if suitable units are employed throughout). Dissolved dyes could be treated similarly.

Owing to the additive nature of α and δ, materials known to be mixtures, such as Brunswick green, lithopone, and dyes laked on substrates, may be treated as single substances, for which values of $[\alpha]$ and $[\delta]$ may be deduced. Even pigment pastes may be so treated, provided that they are made up in the same medium as the paint to which they are added.

Although the experimental work quoted relates to water paints, the principles evolved may be expected to be applicable to other types of paints and to other pigmented systems, such as coloured cements, pigmented paper, plastics, rubber, etc. The data of Amy and collaborators [2] show, indeed, that simplified equations which are based on the same fundamental assumptions as the present accurate formulae are applicable to pigmented papers and lacquers.

In the present investigation attention has been confined to the reflexion of light. Expressions are, however, quoted (viz., (2) and (4)), in which the opacity of films

Table 6. Values of $\Phi[=2R/(1-R)^2]$ corresponding to values of R from 0 to 1

R	·00	·01	·02	·03	·04	·05	·06	·07	·08	·09
0·0	0·000	0·020	0·042	0·064	0·087	0·111	0·136	0·162	0·189	0·217
0·1	0·247	0·278	0·310	0·343	0·379	0·414	0·453	0·493	0·535	0·578
0·2	0·63	0·67	0·72	0·78	0·83	0·89	0·95	1·01	1·08	1·15
0·3	1·23	1·30	1·38	1·47	1·56	1·66	1·76	1·86	1·98	2·10
0·4	2·22	2·36	2·50	2·65	2·81	2·97	3·16	3·35	3·55	3·77
0·5	4·00	4·25	4·51	4·80	5·10	5·44	5·79	6·2	6·6	7·0
0·6	7·5	8·0	8·6	9·2	9·9	10·6	11·4	12·3	13·3	14·4
0·7	15·5	16·7	18·4	20·0	21·9	24·0	26·4	29·1	32·3	35·8
0·8	40·0	44·8	50·7	57·5	66	76	88	103	122	147
0·9	180	224	287	379	522	760	1200	2160	4900	19800
1·0	∞									

is expressed in terms of the same constants. It should, therefore, be possible to utilize the data for predicting the relative hiding powers of paints. No attempt has been made to do this, but there is little doubt that reliable predictions could be made, since opacity measurements by Amy and collaborators on coloured paper gave results in accord with this theory. It can, moreover, be shown mathematically that the present theory leads to the same values for the reflectance of a thin translucent film over a dark background as the Kubelka-Munk theory[4], which has been shown[5] to be applicable to cold-water paints similar to those employed in the present investigation.

REFERENCES

(1) Smith, T. and Guild, J. *Trans. Opt. Soc.* **33**, 73 (1932).
Judd, D. B. *J. Opt. Soc. Amer.*, **23**, 359 (1933).
A detailed explanation of the method of calculating the colour on the C.I.E. basis from the reflexion spectrum is given in the following paper:
Razek, J. *Paint Varn. Prod. Man.* **18**, 332, 368 (1938).
(2) Amy, L. *Rev. Opt.* (*théor. instrum.*), **16**, 81 (1938).
Sannié, C., Amy, L. and Sarraf, J.-M. Ibid. **16**, 86 (1938).
(3) Neugebauer, H. E. I. *Z. tech. Phys.* **18**, 137 (1937).
(4) Kubelka, P. and Munk, F. *Z. tech. Phys.* **12**, 593 (1931).
(5) Judd, D. B. *Bur. Stand. J. Res., Wash.*, **19**, 287 (1937).

DISCUSSION

Dr. R. F. Hanstock. Equation (3) can be identified with one developed by T. Smith (*Trans. Opt. Soc.* **33**, 155 (1931–2)), and both equations can be reduced to the following form:

$$kh = \tfrac{1}{2} \log_e \frac{r(1-Rr)}{r-R}.$$

Amy's constants (a, δ) are related to Smith's constants (k, r) by the equations

$$a = k\left[\frac{1-r}{1+r}\right] \quad \text{and} \quad \delta = k\left[\frac{2r}{1-r^2}\right].$$

Mr. T. Smith. I am particularly pleased to read Dr. Duncan's account of his experiments and to find that he has obtained good agreement between observation and theory. The theory to which he refers partially covers the ground of some work I carried out more than eight years ago. Since then I have attempted on more than one occasion to interest workers in the dyeing and other colour industries in the theoretical results, but without much success. I only recall coming across one worker who appreciated the value of theory in dyeing, and I think he was too engrossed with a theory which I could not accept to be interested in another. The workers to whom Dr. Duncan is indebted for the formulae he uses have established their results by a line of argument entirely different from those I employed; the applications which form the subject of this paper are illustrations only of the simplest set of conditions which may be found. My purpose in drawing attention to this is to state emphatically that the general

structure of the expressions found is imposed by considerations of a very general kind, so that we can claim confidently that they are generally applicable to problems of this type. It happens, however, that in this paper they are not expressed in the form most convenient for numerical use. As tables of hyperbolic functions of suitable accuracy are readily obtainable it is best to use them. The essential formulae are

$$I_0/I = \cosh h\theta + \sinh h\theta \cosh \phi$$

and

$$R/I = \sinh h\theta \sinh \phi.$$

It at once follows, in the author's notation, that for a thin section

$$\alpha + \delta = h_0\theta \cosh \phi, \qquad \delta = h_0\theta \sinh \phi,$$

or

$$\coth \phi = 1 + \alpha/\delta, \qquad h_0\theta = [\alpha(\alpha + 2\delta)]^{\frac{1}{2}}.$$

Since the author only deals with reflected light, it should be noted that he has checked only part of the predictable results.

I might add that the effects of the air-paint surface can be taken into account as a correction to $1/(1-R)$, constant for a given medium.

AUTHOR'S REPLY. It is very satisfactory that the two lines of argument should lead to expressions which are equivalent to one another. The relative convenience of the different expressions depends upon the use to which they are to be put. For a general theoretical treatment, particularly when the transmission of light has to be considered, Mr. Smith's formulae undoubtedly have the advantage, but for the particular problem in which I was interested I feel that the use of $[\alpha]$ and $[\delta]$ is simpler and more likely to be acceptable to the technical worker.

It may be added that the experimental work was not undertaken with the object of testing Amy's theory. It was carried out over ten years ago in testing certain theoretical views of my own, which, however, did not agree well with the experimental results, and I was therefore particularly pleased to find, on the publication of Amy's theoretical views, that they lead to expressions which gave results agreeing so well with my previously determined experimental data.

NORMAL AND ABNORMAL REGION-E IONISATION

By E. V. Appleton, F.R.S., and *R. Naismith, M.I.E.E.
* Radio Department, National Physical Laboratory

Received 29 *March* 1940

ABSTRACT. An account is given of the variation of normal and abnormal region-E ionisation through the sunspot cycle. While the solar ultra-violet light responsible for the production of normal region E varies markedly through the cycle, it is found that such radiation may not be expressed as a single-valued function of mean sunspot numbers, the solar radiation being greater for a given sunspot number during years of rising solar activity than during years when the activity is falling. Normal and abnormal region-E ion-productions have not been found to vary similarly during years of varying solar activity.

§ 1. INTRODUCTION

In a recent paper[1] certain difficulties which arise in the application of the critical-frequency method[2],[3] of measuring upper atmospheric ionisation have been discussed. In particular, it was shown how the presence of abnormal region-E ionisation in the form of clouds or strata, embedded within the normal region E, is indicated by certain features in the curve representing the relation between equivalent height of reflection (h') and electric wave frequency (f). Further evidence relating to the diurnal and seasonal incidence of abnormal region-E ionisation, additional to previous accounts[4],[5], was also given. In the present communication the discussion of the normal and abnormal region-E ionisation is continued, their variations during the sunspot cycle being compared. Further, it is shown how the true critical frequency of normal region E can be estimated even in the presence of abnormal ionisation; while, finally, the heights of maximum ionisation in the two cases are compared.

§ 2. THE INTERPRETATION OF REGION E (h',f) DATA

In the early days of the application of the critical-frequency method of measuring electron concentrations, the distinction between normal and abnormal region-E ionisation was not understood, and the marked variations which are now known to take place in the abnormal ionisation were often erroneously attributed, without qualification, to the normal region E. In 1935[4], however, we presented evidence to show that by recognizing the separate existence and unlike behaviour

of normal and abnormal ionisation it was possible to demonstrate that the normal ionisation followed a very regular law as regards its dependence on the angle of incidence of solar ultra-violet light, and that the great variability of region-E ionisation as a whole was due mainly to the variability of the abnormal region-E ionising agency. In this connection we may usefully recall the characteristics of the (h', f) curve which enable us to make such distinctions between normal, abnormal and intense region-E conditions.

(a) *Normal region-E condition*

When only normal region-E ionisation is present, the (h', f) curve presents the simple form shown in figure 1. The critical frequency of region E for the ordinary

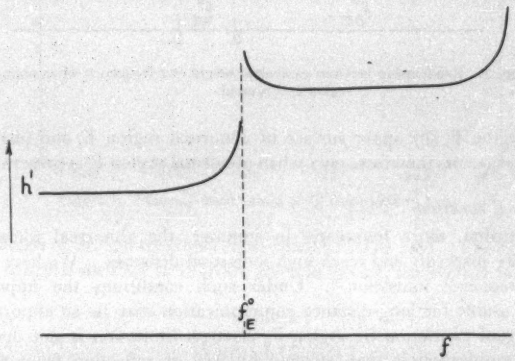


Figure 1. Relationship between equivalent height and frequency (simple conditions).

wave component is then specified by f^{o}_{E} (figure 1). (We concern ourselves in what follows only with the ordinary components.)

(b) *Abnormal region-E condition*

In general the presence of abnormal ionisation is shown by the occurrence of an (h', f) curve of the type given in figure 2. In such a case it is clearly no longer possible to specify f^{o}_{E}, the critical frequency of the normal region E, with precision. International agreement has, however, been reached whereby, under conditions represented by figure 2, the cusp frequency f^{o}_{EE} is taken as the best readily available index of normal region-E critical frequency. As an index of the ionic density of abnormal region-E we take the frequency at which E echoes cease and region-F echoes are simultaneously observed; but if overlap occurs as illustrated in figure 2, we take the mean frequency of the overlap. The penetration of abnormal region E is shown in figure 2 as beginning at (a) and ending at (b), while the critical frequency of abnormal region E is denoted by f^{o}_{AE}. Further evidence of the

presence of the abnormal region-E ionised stratum is the occasional appearance of the so-called M echoes (that is to say, echoes which are reflected, in succession,

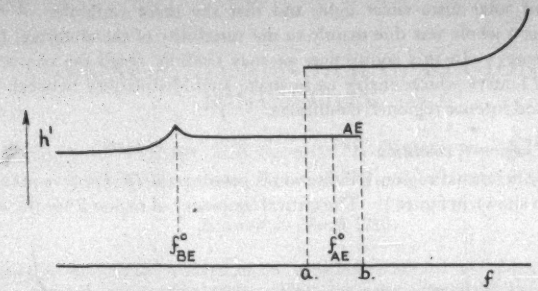


Figure 2. Relationship between equivalent height and frequency when abnormal region E is present.

from (i) region F, (ii) upper surface of abnormal region E, and (iii) region F). Such echoes occur, therefore, only when abnormal region E is present[4].

(c) *Intense E condition*

On occasion, more frequently in summer, the abnormal ionisation may develop very markedly and reach high ionisation densities. We have called this "intense region-E ionisation". Under such conditions the upper limit of frequency usable for long-distance communication may be so abnormally high that the usual restriction by region-F_2 electron limitation is not operative[4].* Such communication is then effected by way of reflection from the intense region E.

§ 3. THE SEASONAL AND SUNSPOT-CYCLE VARIATION OF NORMAL REGION-E IONISATION

The monthly mean values of f^o_E, the critical frequency of normal region E, for the years 1934–9 are plotted in figure 3. The values of f^o_E given here are those noted only when normal region-E conditions prevailed (see figure 1).

(a) *The seasonal variation of normal region-E ionisation*

The applicability of the recombination law in region E was first deduced by us, using the seasonal variation of ionisation at noon. For a latitude λ the ratio of summer to winter noon ionisation (i.e. N_s/N_w) should, according to the recombination law, be equal to

$$\sqrt{\left(\frac{\cos(\lambda-\delta)}{\cos(\lambda+\delta)}\right)},$$

* The examples of abnormal transmission recently described by Pierce[6] can be explained in this way.

when δ is the sun's maximum declination. Since ionisation is proportional to the square of the critical frequency, we can write

$$\frac{N_s}{N_w} = \left(\frac{f_s}{f_w}\right)^2 = \sqrt{\left(\frac{\cos(\lambda-\delta)}{\cos(\lambda+\delta)}\right)} \qquad \ldots\ldots(1).$$

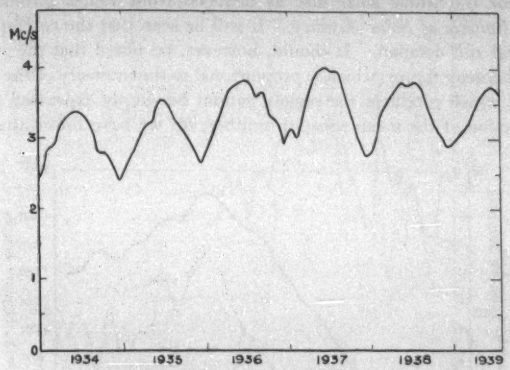


Figure 3. Monthly noon values of f_E° from 1934 to 1939.

For the latitude of Slough this theoretical ratio should be 1·85. The experimental values of $(f_s/f_w)^2$ from 1934 to 1938 are given in table 1, the average being 1·90.

Table 1

Year.	(f_s^2/f_w^2)
1934	1·89
1935	1·96
1936	1·86
1937	1·97
1938	1·80
Mean	1·90

(b) *The variation of normal region-E ionisation through the sunspot cycle*

The general variation of noon region-E ionisation in sympathy with solar activity, previously suspected by us in 1932[8], has been confirmed by the measurements made in succeeding years, leading us to conclude that there is a very marked variation of solar ultra-violet radiation through the sunspot cycle. Since we have now further confirmatory evidence of the applicability of the recombination law we can feel confident in using the expression $(f_E^\circ)^4/\cos\chi$, where χ is the zenith angle of the sun, as an appropriate index of the intensity of such radiation. Using

such an index we can illustrate the variation of ionising radiation with sunspot activity in two ways. The general increase in the index with increase in sunspot numbers is shown in figure 4, where twelve-monthly running means of $(f_E^{\circ})^4/\cos\chi$ are plotted together with twelve-monthly running means of the relative sunspot numbers for the whole solar disc as deduced from values published in the *Quarterly Bulletin of Solar Activity*. It will be seen that the correspondence is both general and detailed. It should, however, be noted that the value of the region-E character figure (which is proportional to the intensity of the solar ultra-violet light which produces the region) cannot be simply expressed as a single-valued function of the mean sunspot number, for we have found that since the

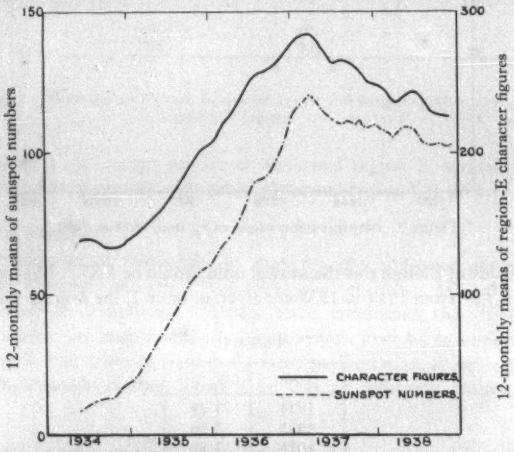


Figure 4. Twelve-monthly running means of region-E character figures and sunspot numbers for 1934–8.

sunspot maximum in the early part of 1937 the region character figure has fallen relatively more rapidly than have the mean sunspot numbers. This remarkable effect is clearly shown in figure 4, where the difference in ordinates becomes progressively less as the cycle advances through the maximum. We therefore conclude that different relations will be required to relate solar ultra-violet emission and mean sunspot number in the two halves, rising and falling, of the sunspot cycle. It is possible that this effect is in some way connected with the fact that sunspots are generally situated in different solar latitudes in the two halves of the cycle.

In figure 5 the actual monthly values of $f_E^{\circ 4}/\cos\chi$ and relative sunspot numbers

are plotted for the same period. A rough correspondence in detail is also apparent. To express this, the departures of the monthly values of both region character

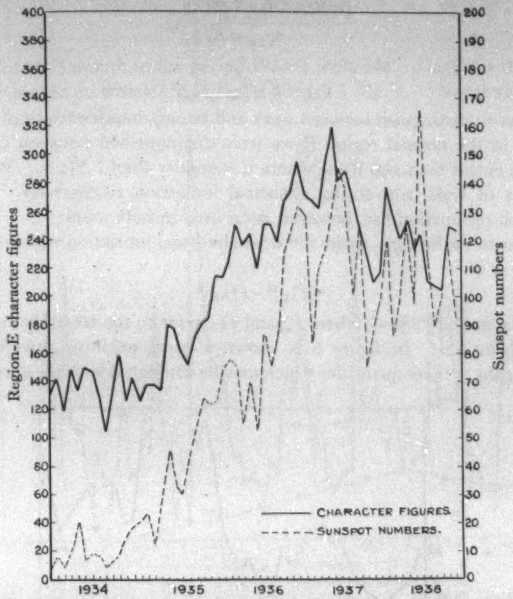


Figure 5. Mean monthly noon values of region-E character figures and sunspot numbers.

figure and sunspot numbers from the mean curves of figure 4 have been extracted and found to have a correlation factor of $+0{\cdot}44$ with a probable error of $\pm 0{\cdot}08$.

§ 4. THE VARIATION OF NOON ABNORMAL REGION-E IONISATION THROUGH THE SUNSPOT CYCLE

Although, as we have previously pointed out, the frequency at which region-F echoes first appear cannot be strictly regarded as the critical frequency of abnormal region E, in that it does not give us an accurate measure of the maximum ionisation in the stratum, we nevertheless believe that f°_{AE} as defined above is a sufficiently reliable indicator of abnormal region-E ionisation-density for us to use it in examining whether the agency causing this ionisation has varied markedly during the sunspot cycle. In making this examination let us assume that the abnormal ionising agency produces ionisation at or near the maximum of region E. Let q_{AE} be the rate of electron production due to it and let q_{BE} be the normal rate of

production of electrons at the same level due to ultra-violet light. Using the familiar formulae we have

$$q_{BE}+q_{AE}=Kf_{AE}^4 \quad \ldots\ldots(2)$$

and

$$q_{BE}=Kf_{BE}^4 \quad \ldots\ldots(3),$$

so that

$$q_{AE}=K(f_{AE}^4-f_{BE}^4) \quad \ldots\ldots(4).$$

In order to distinguish between weak and strong manifestations of embedded ionisation in the normal region E we have distinguished between cases where $(f_{AE}^\circ-f_{BE}^\circ)$ is less than and those where it is greater than 1 Mc./s. We may call these cases of weak and strong abnormal ionisation respectively. The same index of ion production has, however, been used in both cases.

In order to see how q_{AE} varies for weak abnormal ionisation we have examined the variation of

$$(f_{AE}^\circ)^4-(f_{BE}^\circ)^4$$

through the period 1934–8, where f_{AE}° and f_{BE}° refer to the monthly means of the relevant quantities. In figure 6 is shown a curve exhibiting twelve-monthly running means of these quantities which may be compared with the corresponding

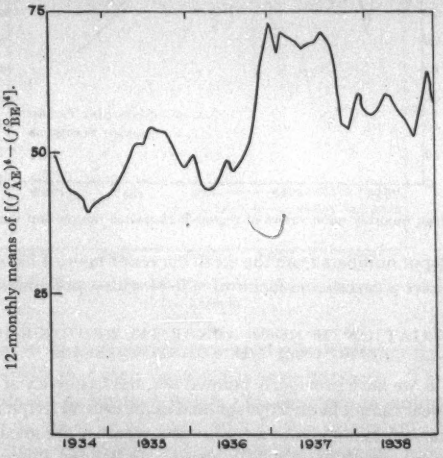


Figure 6. Twelve-monthly running means of $[(f_{AE}^\circ)^4-(f_{BE}^\circ)^4]$ for the years 1934–8.

graph (figure 4) for normal region E. While it is true that q_{AE} during the years 1937–8 is greater than during 1934–5, the relation of this increase to increased sunspot activity may not be significant. In any case it is clear that the increase in q_{AE} is not as large as that established for q_{BE}.

The result of a similar examination for strong abnormal region-E ion-production is shown in figure 7, where the quantity

$$Q = \frac{\Sigma[(f^{\circ}_{\mathrm{AE}})^4 - (f^{\circ}_{\mathrm{BE}})^4]}{n}$$

is plotted for each month in the period 1933–9 inclusive. Here n is the number of days per month on which $(f^{\circ}_{\mathrm{AE}} - f^{\circ}_{\mathrm{BE}})$ exceeded 1 Mc./s. The curve illustrates

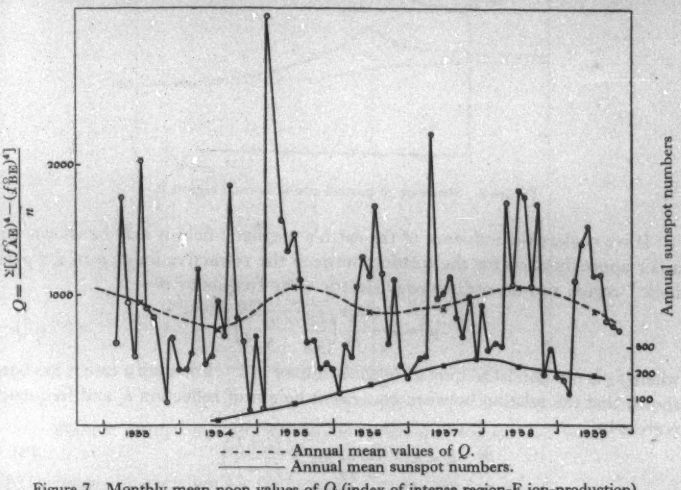


Figure 7. Monthly mean noon values of Q (index of intense region-E ion-production).

the well-known tendency for abnormal ionisation to be most marked in the summer months. On the same graph are shown the annual mean values of Q and the annual mean values of sunspot numbers. It will be seen that during the period 1934–7, when sunspot activity was increasing, there was no abnormally marked increase of abnormal region-E ion-production. It may, however, be noticed that the winter values of abnormal region-E ion-production varied somewhat similarly to the sunspot numbers.

§ 5. THE RELATIVE LEVELS OF NORMAL AND ABNORMAL REGIONS E

It is of interest to draw the theoretical curve exhibiting the relation between equivalent height and frequency for the compound ionospheric structure, consisting of both the normal and abnormal regions. For this purpose we assume region E to be a "parabolic" structure of semi-thickness y_m and ordinary-ray critical frequency f°_{E} and picture the abnormal region E as a more intensely ionised sheet embedded in it. The abnormal-E stratum is shown diagrammatically in

figure 8 at a height y_0 above the lower boundary of the layer, which itself is assumed to be situated at a height h_0 above the ground.

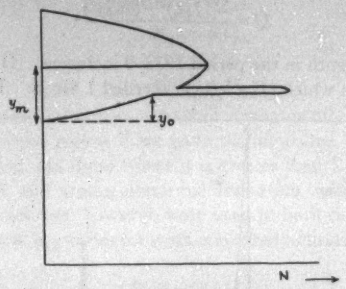


Figure 8. Structure of normal and abnormal region E.

If we neglect the influence of the earth's magnetic field it may be shown that a fair approximation for the relation between the refractive index μ of a "parabolic" region and the ordinary-ray electric wave-frequency is

$$\mu^2 = 1 - \frac{f_E^{\circ 2}}{f^2}\left(\frac{2y}{y_m} - \frac{y^2}{y_m^2}\right) \qquad \ldots\ldots(5),$$

where f_E° is the critical frequency of the ordinary ray. For such a case it has been shown that the relation between equivalent height of reflection h' and frequency is given by

$$h' = h_0 + \frac{y_m}{2}\frac{f}{f_E^\circ}\log_e\left(\frac{f_E^\circ + f}{f_E^\circ - f}\right) \qquad \ldots\ldots(6).$$

Such an equation will apply to that part of the (h', f) curve lying to the left of the cusp frequency f_{BE}° (figure 2). The value of the cusp frequency is evidently that for which the refractive index becomes zero at height y_0. From (5) this is given by

$$f_{BE}^\circ = f_E^\circ\sqrt{\left(\frac{2y_0}{y_m} - \frac{y_0^2}{y_m^2}\right)} \qquad \ldots\ldots(7).$$

For frequencies greater than that of the cusp frequency it may be shown that the corresponding relation between h' and f is

$$h' = h_0 + y_m\frac{f}{f_E^\circ}\log_e\left[\frac{\sqrt{1-\left(\frac{f_E^\circ}{f}\right)^2\left(\frac{2y_0}{y_m} - \frac{y_0^2}{y_m^2}\right)} + \frac{y_0}{y_m}\frac{f_E^\circ}{f} - \frac{f_E^\circ}{f}}{1 - \frac{f_E^\circ}{f}}\right] \qquad \ldots\ldots(8).$$

A typical (h', f) curve obtained from formulae (6) and (8) is shown in figure 9. In constructing this figure the following values of h_0, y_m, y_0 and f_E° have been used:

$h_0 = 100$ km., $y_m = 20$ km., $y_0 = 10$ km. and $f_E^\circ = 4$ Mc./s.

The cusp frequency f_{BE}° is then 3·46 Mc./s. The equivalent height *vs.* frequency relation for values up to this cusp frequency is given by

$$h'(\text{km.}) = 100 + 2{\cdot}5f \log_e\left(\frac{4+f}{4-f}\right) \qquad \ldots\ldots(9).$$

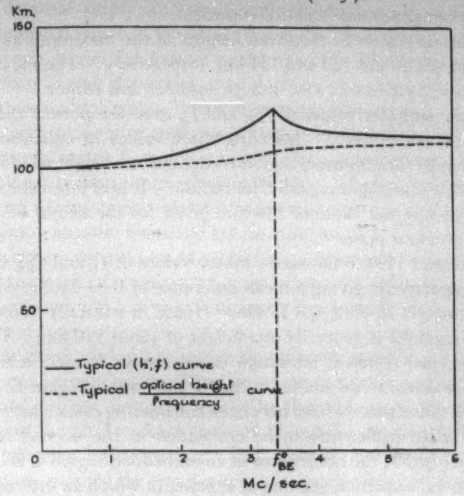


Figure 9. Relationships between equivalent height and frequency, and between optical height and frequency.

The corresponding relation for frequencies greater than f_{BE}° is

$$h'(\text{km.}) = 100 + 5f \log_e\left[\frac{\sqrt{1-\frac{3}{4}\left(\frac{4}{f}\right)^2}-\frac{2}{f}}{\left(1-\frac{4}{f}\right)}\right] \qquad \ldots\ldots(10).$$

At the cusp frequency the equivalent height h' is

$$100 + 10\sqrt{3}\log_e(2+\sqrt{3}) = 122{\cdot}8 \text{ km.}$$

It is also of interest to draw on the same diagram the corresponding curve relating the optical height (which is half the optical path P) and electric wave frequency. In this case, as previously shown, the relation is

$$\frac{P}{2} = h_0 + \frac{y_m}{2}\left[1 - \frac{f_E^{\circ 2}-f^2}{2ff_E^{\circ}}\log\left(\frac{f_E^{\circ}+f}{f_E^{\circ}-f}\right)\right] \qquad \ldots\ldots(11)$$

for frequencies below the cusp frequency f_{BE}°. The corresponding relation for frequencies higher than f_{BE}° is given by

$$\frac{P}{2} = h_0 + \frac{y_m}{2}\left[1 - \left(\frac{y_m - y_0}{y_m}\right)\sqrt{1 - \frac{f_E^{\circ 2}}{f^2}\left(\frac{2y_0}{y_m} - \frac{y_0^2}{y_m^2}\right)} + \frac{f^2 - f_E^{\circ 2}}{f^2}\left(\frac{h'-h_0}{y_m}\right)\right] \qquad \ldots\ldots(12).$$

$(h'-h_0)$ being the equivalent height in the layer. The curve obtained from formulae (11) and (12) after insertion of the values of h_0, y_m, y_0 and f_E^c given above is included in figure 9. When the relative levels are known, it is clear from equation (7) that we can estimate the critical frequency of the normal region E (f_E^c) even when abnormal region E is present.

For the normal region E, the mean heights of the maximum ionisation levels in summer and winter are 120 and 134 km. respectively. The mean value of the semi-thickness y_m is about 22 km. in both summer and winter.

The average summer values of f_{bE}^c and f_E^c over the period 1934–8 are 3·54 and 3·71 Mc./s. respectively. Inserting these values in equation (7), we find $y_0 = 15$ km., that is, the abnormal region is situated at a height of 113 km. or 7 km. below the level of maximum region-E ionisation. It is of interest to compare this with an average summer value of 116 km. given for the height of the abnormal region E in a previous paper [1].

Over the period 1934–8 the mean winter values of f_E^c and f_{bE}^c were 2·67 and 2·63 Mc./s. respectively, giving a mean difference of 0·04 Mc./s. Inserting this value in equation (7) we find $y_0 = 18$ km. Hence in winter it would appear that the abnormal region E is generally at a height of about 130 km. This indicates, then, that abnormal region-E ionisation is not always formed at the same level but exhibits a seasonal trend similar to that of the normal region E.

It should be noted that we have here been considering cases where the abnormal region E has fallen sufficiently to be embedded in the normal region E. As previously described [1], the occurrence of an embedded region E is often the final stage in a temporal sequence, starting at sunrise, in which an intermediate region gradually descends to the normal region-E level. In winter it often happens that, due to the shortness of the day, this process is incomplete at noon. This accounts for the relatively high average noon values of h'_{AE} in winter exhibited previously in a graph showing the seasonal variation.

§ 6. FURTHER DISCUSSION OF THE EXPERIMENTAL RESULTS

The results described above have shown how it is possible to distinguish clearly between normal and abnormal region-E phenomena. The normal region E is found to behave regularly, its maximum ionisation being controlled jointly by the intensity of emission of solar ultra-violet light and the sun's zenith distance. Further confirmation has been obtained of the fact that the electrons in normal region E disappear according to a recombination law, while a rough detailed correspondence of monthly region-E character figures and sunspot figures for the whole disc has been noted. It has previously been shown by us that the ultra-violet light responsible for the formation of normal region E increases by 120 to 150 per cent. during the ascending period of the present sunspot cycle, though it has now been found that such ultra-violet radiation is not a single-valued function of mean sunspot number.

The occurrence of abnormal or sporadic region E, which takes the form of

local ionised clouds or strata embedded in the normal region E, is characterized by great variability in its density and incidence. Our study of the phenomenon has therefore had to be largely statistical in character. Such an examination of the seasonal data has amply confirmed our previous conclusion that both weak and strong abnormal region-E ionisation is subject to some form of solar control. Such ionisation is, in general, greater by day than by night and is stronger during the summer than during the winter, though there is no evidence that it varies as markedly as does normal region-E ionisation with solar activity.

The solar control mentioned above is further exhibited by the variation of the height of abnormal region E, which is found to be just below that of the maximum of normal region E both in summer and in winter. The height of normal region E has been shown to vary most regularly under solar control, the level of maximum ionisation varying steadily from 120 km. in summer to 134 km. in winter.

When, however, we attempt to ascribe the occurrence of abnormal region-E ionisation to any particular agency we meet substantial difficulties. The marked solar control on its origin has been sufficiently emphasized, since it is exhibited in its most marked form during the daytime and in local summer ; yet, since fairly intense ionisation of this kind is frequently observed at night, it is clear that some agency not proceeding from the direction of the sun can also give rise to it.

To proceed further, it would ideally be desirable to compare the results of examinations of the kind we have made at Slough for a series of stations situated at different latitudes. Unfortunately such data are not generally available, and it is necessary to attempt the comparison with incomplete information.

It may first be noted that in our study of the ionosphere at Tromsø, made during the Second International Polar Year[5], we found that the development of abnormal region-E ionisation was often the accompaniment of weak magnetic activity and the precursor of the intense ionisation below region E which gave rise to the "no-echo" condition characteristic of periods of intense magnetic activity. If charged ionising particles are the cause of polar magnetic storms and auroral manifestations it is therefore quite clear that they must similarly be responsible for abnormal region-E ionisation. It will be recalled in this connection that the height of the maximum occurrence of the aurora is approximately that at which abnormal region-E ionisation occurs.

One of the striking features of these polar magnetic storms is their tendency to recur at approximately the same local time for a sequence of evenings. We have noticed the same tendency for abnormal region-E ionisation to develop in this way in temperate latitudes, suggesting the persistence of a localized ionising agency, bearing a fairly constant positional relation to the earth, during a period of a few days' duration. So far as these phenomena are concerned, therefore, abnormal region-E ionisation would be explained in terms of ionising particles which can reach the dark side of the earth.

Our difficulty, however, is to reconcile such an hypothesis with the fact, amply illustrated above, that abnormal and intense E manifestations are subject to marked

solar control. So far as daytime effects are concerned, one might, for example, imagine that the ultra-violet light from the sun had a granular structure, being inhomogeneous over an area corresponding to that of the earth. One must dismiss this, however, for several reasons. The abnormal and intense ionisation tends to occur at a fairly constant level, whereas if the radiation were variable one would expect irregularities to occur at all levels. Moreover, as previously reported, one can watch the gradual development of abnormal region-E ionisation from an intermediate region down to within the normal region E, suggesting a steadily increasing penetrating power. Such a phenomenon is difficult to explain in terms of varying intensity of ultra-violet radiation of constant quality. Also, abnormal ionisation has been found to vary less rapidly than normal ionisation during the sunspot cycle. But perhaps the most decisive information in this connection is that adduced by Berkner and Wells[9], who have shown that at Huancayo, where the sun is nearly vertically overhead during daytime the whole year round, there is very little evidence of abnormal region-E phenomena * compared with that at another southern hemisphere station, Watheroo, at a higher latitude. Such a latitude variation, opposite in character to that exhibited by normal region-E ionisation, would suggest that the normal and abnormal effects are due to agencies of different character.

The fact that abnormal and intense region-E phenomena vary with latitude is further illustrated by the ionospheric observations made by Whatman and Hamilton[10] in North-East Land (lat. 80° N.), who found that maximum effects were observed at approximately 1700 h. local time. At a lower latitude (e.g. Slough, 51° N.) such a maximum is found to occur roughly round local noon.

The foregoing discussion may be summarized as indicating that abnormal region-E ionisation is most probably due to an irregular corpuscular ionising agency which is not markedly effective in equatorial regions but increases in incidence and intensity with increasing latitude. In temperate latitudes it occurs most frequently and most effectively during the day and during the summer. The reconciliation of this solar control with the fact that abnormal E ionisation frequently occurs at night is really the outstanding theoretical difficulty. Either the sun must be a source of the irregular corpuscular ionising agency or else the agency itself must be such that, although of random rate of incidence throughout the day and night, it produces its maximum ionospheric effect when the sun's zenith distance is low. In the latter connection we might picture the collisional detachment of electrons from a bank of negative ions, the concentration of which varies under solar control. In our discussion above we have, for simplicity, assumed that the abnormal and intense region-E electrons are liberated by the

* It is important to note that the abnormal region-E phenomenon is not entirely absent at Huancayo, for Wells states that, at that station, there is no marked critical frequency observed comparable with that at temperate-latitude stations. He describes the transition from region-E to region-F reflection, with increasing frequency, as gradual. Obviously this refers to the abnormal region-E phenomenon, the cusp frequency which indicates the normal region E having escaped notice.

process of ionisation from neutral molecules, but it is possible that a detachment theory of electron liberation would repay further examination.

It is most probable that abnormal ionisation predominates at night in region E, the normal ionisation having dwindled to relatively insignificant proportions. This would explain why the *effective* recombination coefficient of the electrons in region E appears less by day than by night.

Finally, it may be mentioned that if corpuscles are responsible for region-E ionisation at levels of 110 to 130 km. it is easy to show that such corpuscles must have a penetrating power of about 1 cm. in air at normal temperature and pressure.

§ 7. ACKNOWLEDGEMENT

We acknowledge gratefully the assistance of Dr. W. J. G. Beynon in this work, which was carried out as part of the programme of the Radio Research Board of the Department of Scientific and Industrial Research.

REFERENCES

(1) Appleton, Naismith and Ingram. *Proc. Phys. Soc.* **51**, 81 (1938).
(2) Appleton. *Nature, Lond.*, **120**, 197 (1931).
(3) Appleton and Naismith. *Proc. Roy. Soc.* A, **137** 36, (1932).
(4) Appleton and Naismith. *Proc. Roy. Soc.* A, **150**, 685 (1935).
(5) Appleton, Naismith and Ingram. *Philos. Trans.* A, **236**, 191 (1937).
(6) Pierce. *Proc. Inst. Radio Engrs, N.Y.*, **26**, 892 (1938).
(7) Appleton. *Proc. Roy. Soc.* A, **162**, 451 (1937).
(8) Appleton and Naismith. *Proc. Phys. Soc.* **45**, 248 (1933).
(9) Berkner and Wells. *Terr. Mag.* p. 73 (March 1937).
(10) Whatman and Hamilton. *Proc. Phys. Soc.* **50**, 217 (1938).

ERRATUM

Page 251 (vol. 52, part 2), line 9 from end (discussion by J. W. Perry of paper by L. J. Comrie) : for *spherical* read *aspherical*.

OBITUARY NOTICES

RALPH ALLEN SAMPSON

THE life of Ralph Allen Sampson, who was born on 25 June 1866 and died on 7 November 1939 at Bath, divides quite naturally into three periods. There is the early formative period, then the Durham period, and finally the Edinburgh period.

Of Cornish descent, Sampson was born in County Cork, but at an early age was taken with his family to Liverpool. There his education was delayed by his father's financial circumstances, but at the Liverpool Institute he made such rapid progress in five years that he entered St. John's College, Cambridge, as sizar, gained a scholarship and became third wrangler in 1888. Next year he gained the first Smith's Prize and was afterwards elected to a fellowship at his own college. In 1891, having spent the interval as mathematical lecturer at King's College, London, he returned to Cambridge as the first Isaac Newton Student. As an outcome of his work in the next two years Sampson produced a substantial memoir "On the rotation and mechanical state of the sun" (*Memoirs R.A.S.*) in which the hypothesis of radiative equilibrium is proposed as an alternative to the convective equilibrium hitherto assumed. This choice of subject illustrates his own natural leaning towards astrophysics, but the strongest personal influence which affected his career from these years at Cambridge came from his contact with J. C. Adams.

In 1893, at the early age of twenty-seven, Sampson went north and became professor of mathematics, first for two years at the Durham College of Science at Newcastle and after that at Durham University. For a man with his capacity for work the academic duties of his office were too light to absorb his energies, and he served astronomy as though that had been the only occupation of his life. This activity was shown in three main phases. On the death of J. C. Adams, Sampson was chosen as editor of his astronomical and mathematical work in a collected form. The orderly arrangement of the unpublished papers proved a difficult and laborious task in which Sampson evinced a distinct gift of patient scholarship. He also provided a memoir in which the circumstances attending the discovery of Neptune are critically reconstructed. Sampson would have made an ideal editor for an edition of Newton's works and the fact was not overlooked, but no project of this kind went further than suggestion during his life.

Durham possesses a small observatory and Sampson naturally assumed the direction of it. Funds being available for the purchase of an instrument, an almucantar of novel design was added to the equipment. Owing to the irregular effect of temperature changes the results obtained proved disappointing. The

fault lay probably in the design of the instrument, certainly not in any lack of vigour in prosecuting the programme of observations. Sampson's readiness to adopt a type of instrument of which little was known was an example of his courage and originality.

But the principal work performed in the Durham period, and the chief work of his life, was devoted to the four great satellites of Jupiter. This fell into three parts, the introduction of fresh observations, the development of a revised dynamical theory and the preparation of new tables. The observations previously used had been chiefly eclipses of the satellites and were known to be subject to strong systematic errors. Pickering at Harvard had initiated the photometric observation of the satellites instead of recording merely the times of disappearance and reappearance, thus substituting a curve for comparison in the place of a doubtful point. This gave new material which Sampson reduced and used. The theory, which is subject to the relation $3n_2 = 2n_3 + n_1$ between the mean motions of the three inner satellites, was virtually completed at Durham, but was published (*Memoirs R.A.S.*) later, in 1921, after receiving some final additions. But the tables based on this theory were published in 1910, and they are now used for the calculations of the national ephemerides. Observations with the heliometer and by photography at the Cape Observatory have caused some doubt as to the sufficiency of the Harvard photometric results, but Sampson carried out an immense task and his work marks a necessary stage in the advance towards a more perfect representation of the Jovian system.

Though he had never served a regular apprenticeship in an astronomical observatory, there could be no surprise when Sampson, with so much astronomical work to his credit, was appointed Astronomer Royal for Scotland in 1910. Like others he was to experience the difficulties imposed by the war years, and towards the end he suffered from failing health which led to his retirement in 1937. Ten years earlier he had gained official sanction for much-needed additions to the Observatory, and in 1932 a 36-inch reflector and a 10-inch photographic instrument fitted with a Cooke triplet lens were installed. The prospect of obtaining the reflector may have led his thoughts to the study of the properties of reflecting combinations, on which he had published papers as early as 1913. His interest in optics had always been far from superficial (a word utterly alien to Sampson's character) and his selection as an honorary member of the Optical Society was fitting.

At Edinburgh, Sampson completed the measurement of the zones of the Perth (W.A.) zones of the Astrographic Catalogue undertaken by the Observatory ; the results are unfortunately still unpublished. He never returned to the heavy work in dynamical astronomy which had occupied so much of his energies at Durham, but turned rather to the physical side which had a natural attraction for him. In the years before the war the dissemination of wireless time-signals had given the clock an importance even greater than had been previously realized, and Sampson gave a critical study to its performance. After the war Mr. W. H.

Shortt was developing improvements which culminated in the Shortt free-pendulum clock, now a standard observatory equipment. In this work Sampson co-operated from the astronomical side and shares in the credit due to a most important practical advance, one consequence of which is that the time-signals from different stations are now more accordant than they were formerly.

It was doubtless Sampson's view, on which he certainly acted, that it is a mistake for an observatory with a small staff and academic responsibilities to become immersed in a heavy programme of routine work, and that it is far more profitable to seize the opportunity to develop new methods. The introduction of photo-electric methods gave him such a chance. His first idea had been to apply the method directly to the telescopic image as others were doing. The next step was to substitute the photographic image, and this led to a careful study of the properties of the plate as a photometric medium. Finally, when the conditions had become fairly familiar to himself and his assistant E. A. Baker, the photographic spectra of stars were made the subject of investigation and the first steps were taken in the determination of colour temperatures. Refinements and further developments have been introduced into the technique by others, but the essentials of a new method were worked out at Edinburgh. One result of the research was to show the necessity for the improved equipment of the observatory, which has been met to the extent already mentioned.

It is tempting to compare the career of Sampson with that of his great contemporary E. W. Brown who died in the previous year. Both were products of the same school, Cambridge. Both became professors of mathematics at an early age, Sampson at Durham, Brown in America. Both found an outlet for their energies in dynamical astronomy. Brown devoted himself to the lunar theory, and it is his tables which are used today. Sampson took up the problem of Jupiter's satellites and his tables are now in use. The complex details of the lunar theory require the devotion of a life's work. Brown was happy in seeing his task finished, and gave his remaining years chiefly to an examination of some of the finer points in planetary theory. Sampson on the other hand turned away from what had been his main occupation, assumed the direction of an important observatory, indulged his taste for astrophysical research and discharged the ordinary duties of a professor as long as health allowed. Thus Brown's life displays a more concentrated unity, Sampson's a greater versatility and a spirit of adventure which rarely finds scope in a scientific life. Both men received the Gold Medal of the Royal Astronomical Society, Brown in 1907, Sampson in 1928. It would be improper to press a comparison of this kind too far, but one common characteristic cannot fail to be noted, the dauntless capacity for hard work. Finally, while one died a bachelor, Sampson leaves a widow, a son and four daughters, all of whom survive him.

H. C. P.

CLEMENT OSBORN BARTRUM

CLEMENT OSBORN BARTRUM died on 28 April 1939. He was born in Bradford on 6 October 1867 but lived all but the first nine years of his life in Hampstead.

He never held any scientific appointment, but his interest in scientific matters was throughout his life keen and extensive. He was elected a fellow of the Physical Society in 1898 and later became associated with the Royal Astronomical Society and the British Astronomical Association. Of the last he was a secretary for the last nine years of his life, involving work which gave happiness both to himself and his fellow members. He was a founder of the Hampstead Scientific Society, thirty-six years ago, acting for twenty-six years as secretary and for ten as treasurer.

Of his many scientific interests the principal was probably horology. In 1913 he devised and set going a slave pendulum which timed the impulses maintaining the swing of the master pendulum and which was in turn synchronized to the beat of the master. This was the forerunner of the well-known Shortt clock and one of few schemes that really worked. A difficulty in synchronizing is "hunting": the slave goes alternately too fast and too slow. This he overcame by a most ingenious scheme which he called "negative backlash", thus securing not only synchronization in *phase* but also in *rate*. In this the method is superior to that of the Shortt clock. The Science Museum has secured the original model on loan as a going concern.

This invention was described in a paper to the Physical Society on 15 February 1917 entitled "A Clock of Precision". All interested in the subject admire the ingenuity of the scheme as well as the manner in which the device and its theory are expounded.

Later Bartrum fitted a pendulum with a compensator for changes in barometric pressure as an alternative to using an airtight case.

He leaves a widow, a son and two daughters. These all shared with him a love of music as performers as well as listeners. E. C. A.

REVIEWS OF BOOKS

Electromagnetics, by ALFRED O'RAHILLY. Pp. xii + 884. (London: Longmans, Green and Co., Ltd., 1938.) 42*s*.

In "Electromagnetics" Professor O'Rahilly makes the startling claim that a theory proposed by Ritz as long ago as 1908, yet overlooked and hitherto little known even within the scientific world, explains electromagnetic phenomena more satisfactorily than the combined theories of Maxwell, Lorentz and Einstein, theories nowadays adopted in their main lines by practically all mathematical physicists.

As the challenge thereby launched constitutes by far the most important aspect of the work, it seems appropriate to attend in the first place to the sections of the book devoted to it, leaving for later consideration various other points raised, which also are in the nature of challenges—for challenge seems to be a favourite sport with the author,—although they are not likely to arouse as much interest as the Ritz versus Maxwell-Lorentz-Einstein controversy.

An introduction to vector analysis, potential and free energy of singlets and doublets, much along the usual lines, is followed by the derivation of a formula for the force between moving charges, the so-called Liénard-Schwarzschild force-formula, to which constant reference is made subsequently, and to which the greatest importance is attached: "We have completely eliminated the magnetic and electric vectors, and we gave the law of elementary action as the complete and adequate representation of the current electron theory. The field in the aether does not enter at all; to obtain the field anywhere we must put an electric charge there. It would be otherwise, of course, if, as Maxwell supposed, **E** and **H** modified or moved the aether. But in reality we observe only **F**, the force on unit charge . . . ."

Considerable space and great pains are devoted to a demonstration of the equivalence between this force-formula and Maxwell's equations, as modified by electron theory, and it is claimed that all writers who make use of the latter accept the former, although in most cases unwittingly. The contention that formulae which yield directly those quantities observed in experiments are superior to those giving supposed states of an imaginary medium may have much to commend it, but since in the present case the avowed object is the comparison of a new theory with generally accepted theory, it would perhaps have been better to have expounded the latter in the usual way, which has over the alternative treatment, as Professor O'Rahilly himself admits, the advantage of mathematical elegance and ease of manipulation.

Only after 500 pages, making up more than half the book, is the reader introduced to the rival theory sponsored by the author: "Ritz's initial assumption is that each electrified point emits, at each instant and in all directions, fictitious infinitely small particles, all animated with the same radial velocity c relative to the origin; so that the aggregate of the particles emitted at the instant t' by a moving electron S forms at any subsequent instant t a sphere of radius

$$\rho = c(t - t').$$

The principle of relativity of motion in its classical form, says Ritz, requires that the waves emitted by a system in uniform motion, not subjected to any external sensible

influence, should move with this system so that the centre of each spherical wave continues to coincide with the electron which has emitted it and the radial velocity is constant and equal to c."

The formula for the force between moving charges deduced from Ritz's assumptions, is shown to be consistent with experimental results on the force between closed currents, the scattering of α particles, the forces between electric currents and electric charges in motion, etc. The great merit of Ritz's formula lies in its complete agreement with the principle of relativity: the force between a moving charge and a moving current-carrying circuit depends only on the relative velocities of the charge and the circuit, not on their absolute velocities. Thus no force is to be expected between a charge and a current when both are stationary on the earth. This easily verifiable result is reached much more simply by Ritz's theory then by Lorentz's electron theory, which meets the experimental fact of absence of force by postulating a compensating distribution of positive and negative charges on the wire owing to its translation through the aether. In the same way Ritz, whose formula predicts that the force between a magnet and an electric charge depends only on their relative velocity, explains experimental results in a more satisfactory manner than Einstein, who for the purpose finds himself forced to call to his aid the fairly complicated relativity transformation of Maxwell's equations. Finally, in contrast to the theory of Lorentz, that of Ritz conforms with the extremely simple and satisfactory hypothesis of a point-charge electron of constant mass.

An instructive section is devoted to a number of classical experiments in electromagnetism, of which that of M. and H. A. Wilson (magnetic dielectric rotating in magnetic field) is among the most interesting, but the solutions and results are not capable of deciding between Ritz's and the classical theory. Moreover, it has been shown by Howe that the majority of the results in question can be obtained by elementary methods, which do not necessitate the use of complicated albeit powerful tools like Maxwell's equations, the formulae of Liénard or Ritz, or the theory of relativity. All the same one feels inclined to agree with the comment: "We have tested Ritz's formula against experimental results, and we have found that it stands the test better than that of Liénard."

The introduction of Ritz's theory and its comparison with accepted theory as a means of explaining experimental results, occupy 131 pages, or about a seventh part of the book, and it could have been made much shorter still. As to the rest, the very beginning is fairly orthodox up to the proof which purports to show that the law of force in magnetic media, $f=mm'/\mu r^2$, "universally given in our text-books", is incompatible with the existence of permanent magnets, and therefore incorrect. To "illustrate the point by a simple particular case" the author applies this formula to the calculation of magnetic intensity outside a uniformly magnetized sphere in two cases: (*a*) if the sphere were "surrounded by an infinite medium of permeability μ_0", (*b*) if it "existed in an unpolarized medium", and the well known result that the ratio of intensities differs from $1/\mu_0$ is held to invalidate the law of force $f=mm'/\mu r^2$. Unfortunately the reasons in support of this point of view are not made clear.

It is conceivable that highly accurate measurements on a system of bulky magnets transferred from a fluid of permeability μ_0 to a fluid of permeability μ would yield a ratio of forces different from μ/μ_0, but the reason for the difference is not that the formula $f=mm'/\mu r^2$ is incorrect, but rather that, since the magnets themselves occupy a portion of the space, the permeability is not μ_0 or μ everywhere, as postulated by the formula, which applies to poles, or to "needles" of lateral dimensions negligible compared with their lengths.

The necessity for introducing a " magnetic constant ", β, and writing the formula in vacuum $f=mm'/\beta r^2$ is stressed throughout the book, and those who have been in the habit of writing $f=mm'/\mu_0 r^2$ and calling μ_0 " permeability of vacuum " become the butt of trenchant, if amusing criticism. Yet although much space and not a little " discourse," both at the beginning and at the end of the treatise, are devoted to a justification of this proposed deviation from current practice, the reader is left with the impression that the quarrel is one of mere words and definitions: some writers choose to call " permeability " what the author insists on calling " magnetic constant," and advocate the term " relative permeability " for what the author calls " permeability ", but the formulae lead to identical results. The context indicates (e.g. see p. 610) that Professor O'Rahilly would have no objection to the statement that the permeability of air at some particular pressure is 1·000 000 4: exhaust gradually: the permeability decreases and tends towards unity. Some writers find it convenient to say that in the process of passing from air to vacuum the permeability has changed from 1·000 000 4 to 1, and this nomenclature seems logical to many besides the " uneducated " of p. 40. In the usual acceptation of the terms, the property which vacuum lacks is susceptibility, not permeability; susceptibility is nil for vacuum, very small for air, fairly large for nickel and iron.

When the time comes for assigning a value (p. 65) to the " magnetic constant β ", the choice falls on unity, in conformity with generally accepted practice. This choice anyway is a matter of convenience, and the author will not be blamed for following in the footsteps of so many orthodox scientists, but it is misleading to state that " there is no reason, theoretical and practical, for doing otherwise." It seems a pity for an opinion on the matter to be presented as unchallengeable truth, especially when the author, to judge from a number of quotations in the same section, has intimate knowledge, if not always correct understanding, of the work of Giorgi, who has given very good reasons, theoretical and practical, for doing otherwise.

Quotations are frequently criticized which yet are perfectly reasonable when taken with their context. Thus in p. 82: " It is quite incorrect to say with Planck that 'one can ascribe to the vacuum any arbitrary value of the dielectric constant, as is indicated by the various systems of units'. Electrical units are varied by changing α, not by varying κ_0 ". But the point is that Planck, and most other scientists too, denote by κ_0 precisely what Professor O'Rahilly, and very few, if any, others choose to denote by $\alpha\kappa_0$. Planck's statement is in perfect agreement with his definitions. The general tenour of the assertions made in connection with these electric and magnetic constants α and β calls to mind the remark of p. 392: " Conclusions in physics are never quite so certain as up-to-date dogmatists imply."

Over 250 pages are devoted to units and dimensions. With words coined to replace those in current use, the author proceeds, " mostly by way of negative criticism," to justify his " elementary intelligible account of a subject which has not only become a bugbear to students but has misled international congresses of experts into talking nonsense ". Unfortunately the lengthiness of this justification is likely to hinder the propagation of the views therein expounded. Why, for instance, so much " discourse " to prove that a velocity of 88 feet per second is the same as a velocity of 60 miles per hour? After a lengthy " explanation " of a short quotation from Sommerfeld (p. 807) comes the claim: " We have therefore ousted all the meta-physical mysteries. We have not only vindicated common sense, we have also provided all the proposed formulae with simple meanings accessible to any schoolboy, entirely too elementary to be the subject of discussions and votes at learned international congresses." In view of the length and complication of the preceding discussion, the general feeling will probably be that this claim is highly exaggerated.

It is nevertheless sincerely to be hoped that criticism of parts of the work will not deter from its serious study, for the arguments in favour of Ritz's theory seem well founded, and at any rate no physicist can ever regret having read such a stimulating, entertaining and unconventional treatise. A sequel is promised, dealing more specifically with those mechanical and optical phenomena which fall at present within the province of relativity. Presumably it offers explanations on some other, perhaps more satisfactory hypothesis, of the advance of perihelion, the deflection of light, the displacement of Fraunhofer lines, and kindred phenomena, which are barely mentioned (p. 544) in the present volume. A presentation concise, elegant and attractive, of the kind, in fact, at which some of our "physics-popularizers" have proved such adepts, and avoidance of digressions on relatively unimportant subjects like Eddington's "elephant sliding down a grassy hillside" (p. 771), the "dissociation of Jeans' dual personality" (p. 645), the "followers of Einstein fishing in troubled waters" (p. 601), may well win for Professor O'Rahilly the satisfaction of seeing his "purely scientific challenge either answered or accepted".

P. V.

The Psychology of Physics, by BLAMEY STEVENS. Pp. xvi+278. (Manchester: Sherratt and Hughes, 1939.) 7*s*. 6*d*. net.

The nature of this book is best conveyed by a few quotations from the preface.

"The difference between intuition and psychology is that the former uses instinctive knowledge subconsciously, and the latter states this knowledge in so many words, that is to say as postulates of perception, and then reasons from these postulates."

"I therefore postulate the instinctive character of space, time and inertia, and deduce the physical laws of nature from the postulates .... The result of these deductions is so satisfactory that the truth of the postulates can hardly be doubted."

"The result of the development of the perceptual theory is iconoclastic; one after another the main axiomatic props of modern physics are shown to be fallacious, namely ... the equivalence of mass and total energy, ... the Doppler effect theory, ... the electric divergence law ..."

"My criticism is ... equally constructive ... I thus deduce Newton's mechanics and Maxwell's field laws from the psychological postulates."

"A table of physical constants with their symbols and logarithms is added."

Only one criticism need be added. If Mr. Stevens's object is to produce a text-book for the "young pupils" whom he "expects to instruct", he is unwise to omit all descriptions of facts and thus to force his readers to derive from works steeped in the errors that he attacks all their knowledge of the meaning of technical terms (such as Doppler effect) that he uses so freely in his arguments.

N. R. C.

Probability and Frequency, by H. C. PLUMMER, M.A., F.R.S. Pp. xi+277. (London: Macmillan and Co., Ltd., 1940.) 15*s*.

Professor Plummer has produced what he himself describes as "an approach to probability and statistics from a simple mathematical point of view". In doing this he does not claim that a knowledge of the mathematical principles is sufficient for an understanding of statistical methods in their manifold fields of application. On the contrary, "the abstract science of statistics" is rightly considered to be "only a means to an end, not an object in itself. Its results can be judged, and useful deductions can be drawn from them, only in the light of expert knowledge of the material on which they are

based". The aim of the book, therefore, is to teach experimenters how to handle mathematical statistics as a useful tool in their researches.

Without entering into philosophical controversies, probability is defined as a practical hypothesis, the usefulness of which is to be proved by comparing its logical consequences with observational facts. This is immediately shown with the help of examples on the "probability of discrete events," which is the subject of the first chapter. Simple examples on permutations and combinations (which serve to show the dependence of results on the precise definition of "equally likely") are followed by problems requiring a gradually increasing amount of mathematical technique. General principles are introduced by examples, so that the laws of addition and multiplication, Bernoulli's theorem and Bayes's theorem appear as natural generalizations of particular problems. In the light of modern experience the importance of Bayes's theorem has perhaps been overstressed. Of considerable interest is the problem of matching cards (pp. 20–21 and 38–39), which has found recent application in psychological research.

A chapter on continuous probability is followed by the "theory of errors", giving an account of the classical work on the normal curve of errors and discussing difficulties in the application of "normal theory" to the rejection of extreme observations. It would have been interesting to include here an account of modern work on small sample theory, which solves some, at any rate, of the difficulties. The method of least squares is introduced in terms of an astronomical problem that leads to linear equations for the small corrections. This is certainly the best way of introducing the linear hypothesis to physicists. The general solution is then given in terms of normal equations and determinants and is illustrated by a numerical example.

The chapter on statistical distributions describes the conceptions of elementary statistics such as histograms, frequency polygons, measures of location and scale, moments and their calculation, Sheppard's correction, etc. Fitting by the method of moments is illustrated with the help of a harmonic analysis, which tends to over-simplify the problem. The calculation of parameter-estimates from moments is shown by using an aggregate of two normal curves with two different means and standard deviations. This example is of importance when analysing the cause of an observed non-normality.

The last chapter, on correlations, deals with distribution problems in several variables. Beginning with classical conceptions such as the multi-normal distribution, the correlation coefficient and ratio, etc., concise mathematical proofs are given for the random sample distributions of a number of recently developed statistics. The technique used is integration in hyper-space, which is straightforward in principle and, although clumsy in mathematical appearance, is perhaps most easily understood by the student. More advanced methods, such as the use of moment-generating functions, have been, it seems, purposely omitted.

Two short statistical tables are appended, based on Legendre's table of $\log_{10}\Gamma(x)$ and Burgess's table of $\operatorname{erf}(t)$. It is gratifying to note that the errors in the latter have been eliminated.

This publication is a mathematical textbook mainly concerned with problems of distribution. It will, therefore, be of great help to the beginner who desires mathematical rigour but would find it difficult to consult the original papers to which he is referred by most of the existing textbooks on statistical methods. H. O. H.